

Exploring the Potential Role of Staff Surveys in School Leader Evaluation

Whitney Kozakowski, Brian Gill, and Menbere Shiferaw

August 2021



This report was funded by the U.S. Department of Education's Institute of Education Sciences (IES) under contract ED-IES-17-C-0006 by Regional Educational Laboratory Mid-Atlantic administered by Mathematica. The content does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Exploring the Potential Role of Staff Surveys in School Leader Evaluation

Whitney Kozakowski, Brian Gill, and Menbere Shiferaw

August 2021

The Mid-Atlantic Regional Educational Laboratory partnered with the District of Columbia Public Schools (DCPS) to explore the potential use of teacher surveys in school leader evaluation. The DCPS evaluation system, like many others, currently consists of two components: an assessment on how well a school performs on a set of student achievement metrics (such as proficiency on standardized tests) and an assessment by a supervisor of the principal's leadership across multiple domains. Incorporating teacher surveys could provide an additional perspective on principals' leadership and performance. Examining data from two teacher surveys that DCPS has used (Panorama and Insight), the study found that:

- Most of the domains and scales from the teacher surveys can meaningfully differentiate among schools.
- School averages of (Insight) survey domains are more similar for adjacent years in which a school had the same leader than when the school leader changed, suggesting principals have effects on these measures.
- When a school gets a new principal, (Insight) survey domain scores change more than student proficiency rates, suggesting (consistent with theory) that principals may affect teachers more immediately than they affect student outcomes.
- All but two of the teacher survey domains and scales have small or moderate correlations with supervisors' ratings of principals' leadership, with values ranging from 0.26 to 0.46. Teacher survey measures are most correlated with three of the domains assessed by DCPS principal supervisors: Instruction, Talent, and School Culture.

Overall, our findings suggest that it could be useful for DCPS to use elements of teacher surveys to bring in teachers' perspectives on principals' leadership related to instruction, talent, and school culture. Other districts may also wish to consider employing teacher surveys to gain an additional perspective on principals from staff who interact with the principal every day.

Why this study?

The District of Columbia Public Schools (DCPS) system uses a multi-measure system called School Leader IMPACT to evaluate its school leaders. In 2019, DCPS partnered with the Mid-Atlantic Regional Educational Laboratory (REL) to examine and enhance the reliability and validity of this system. The REL began by studying the variability of the components of the School Leader IMPACT system to better understand the role of random variation in the student test score component. One recommendation that came from the coaching project was that incorporating an additional measure into the evaluation system might improve the validity and reliability of School Leader IMPACT and potentially provide useful feedback to leaders (see appendix A). This study examines data from existing staff surveys to assess whether and how they might inform judgments of school leaders' leadership and performance.

Evaluating school leaders is difficult because the job of a principal is so varied from day to day and hour to hour. Measuring school leaders' impacts on student outcomes is also challenging, because school leaders do not teach students directly and are likely to influence outcomes over a multi-year period (Chiang et al., 2016; Coelli & Green, 2012; Dhuey & Smith, 2014; Grissom et al., 2015). School leaders' indirect influence on students, through channels such as school climate, teacher turnover, and teacher instruction, has been documented in the literature

(Burkhauser, 2017; Leithwood et al., 2004; Sebastian & Allensworth, 2012). Incorporating staff input could provide useful information on principals' leadership and performance, thereby capturing some of their indirect influence on students from the perspective of those who work with them every day.

Existing research suggests that staff ratings of principals are associated with student achievement gains and have potential to provide useful feedback to leaders. In one study, ratings of principals produced by teacher surveys were significantly related to school value-added scores, suggesting principals' behaviors affect student achievement (Liu et al., 2014). In a rigorous two-year study in eight school districts of a comprehensive evaluation system that included school leader evaluations informed by staff surveys, the authors found that implementing the staff-informed leader evaluation system alongside a teacher evaluation and feedback system produced significant improvements in principal leadership (instructional leadership and teacher-principal trust) and increased student achievement in math (but not reading) (Garet et al., 2017). The study also yielded evaluation scores that were distributed across the entire range of possible scores (indicating their potential to signal a need for improvement), producing more differentiation than many current evaluation systems where scores cluster near the top of the scale.

Key steps in understanding whether staff input could be a valuable addition to DCPS's school leader evaluations include understanding: whether the surveys capture meaningful differences between schools; how much these differences might be due to principals (as opposed to other differences between schools); and the relationships between existing school leader evaluation metrics and staff surveys. This study addresses these questions. As a result, this study can inform DCPS's decision whether or not to include feedback from staff surveys into the School Leader IMPACT system. Moreover, findings will provide additional evidence for other school districts on whether staff input can improve school leader evaluation systems.

Box 1. Key terms

District of Columbia Public Schools School Leader IMPACT scores: The DCPS School Leader IMPACT evaluation system measures school leaders' effectiveness with the goals of setting clear expectations and providing feedback for improvement. Under the School Leader IMPACT system, school leaders receive a final evaluation score, which is a summary measure of their leadership and performance. The final score is composed of two equally weighted parts: Leadership Framework standards and Student Outcome Goals (described below).

Leadership Framework standards: The Leadership Framework aims to assess the effectiveness of a school leader's practices with respect to improving student learning, as gauged by the school leader's supervisor. The Leadership Framework assesses school leaders relative to six standards using both qualitative and quantitative measures: Instruction, Talent, School Culture, Operations, Family & Community, and Personal Leadership (see appendix B).

Student Outcome Goals: Student Outcome Goals measure school-level learning outcomes for students. There are two components: a student achievement goal component, focused on student proficiency in the standardized test used for accountability purposes, and a school-specific goals component, focused on other high-need areas for a school's success which could include other student achievement measures or school culture.

Insight Survey domains: The District of Columbia Public Schools (DCPS) partnered with The New Teacher Project to administer the Instructional Culture Insight survey to teachers to measure their perceptions of school instructional culture. The survey consists of items grouped into the following 11 domains (those marked with * refer to domains that were added in 2017/18):

- **Academic Opportunity:** Perceptions of whether rigorous academic practices are visible in classrooms and whether school leaders have the necessary resources and knowledge to support these practices.
- **Diversity, Equality, and Inclusion*:** Perceptions of whether diversity and inclusion are a clear priority at their school and whether the school's talent management practices and overall culture advance that priority.
- **Evaluation:** Perceptions of how clearly performance expectations have been communicated, how much teachers agree with those expectations, and how accurately evaluations represent teacher performance in this school.
- **Family & Community Engagement*:** Perceptions of the relationships teachers have built with students' families and the systems in place for keeping families involved in students' education.

- *Instructional Planning for Student Growth*: Perceptions of whether teachers share student assessments or receive instructional planning support to make sense of student assessment data and use it to improve their instruction.
- *Leadership*: Perceptions of the effectiveness of school leaders (including the vision they set for the school, the extent to which they communicate and follow through on clear goals and priorities) and whether they seek upward feedback from their teachers.
- *Learning Environment*: Perceptions of the consistency of expectations and consequences for student conduct, as well as leadership support in maintaining a productive learning environment.
- *Observation and Feedback*: Perceptions of how frequently teachers are observed and how consistently they get helpful feedback to improve their instruction.
- *Peer Culture*: Perceptions of colleagues within the school, including whether peers share similar priorities and provide examples of exemplary teaching.
- *Professional Development*: Perceptions of experiences with professional development and how well it has helped teachers learn new, relevant skills and improve their instructional practice.
- *School Operations*: Perceptions of the efficiency and effectiveness of daily operations in the school building, including non-academic services, student information systems, and overall cleanliness.

The Insight survey also includes the index score, which is a summary measure based on teacher agreement with three indicators of instructional culture: that the expectations of effective teaching are clearly defined, that there is a common vision of what effective teaching looks like, and that their school is committed to improving instructional practice.

Panorama teacher survey scales: DCPS partnered with Panorama Education to administer the Panorama survey to instructional staff to measure their perspectives on social-emotional learning (SEL), school climate, and engagement. The Panorama teacher survey consists of items grouped into the following five survey scales measuring teachers' perceptions of the following:

- *Perseverance*: How well their students can persevere through setbacks to achieve important long-term goals.
- *Rigorous Expectations*: The degree to which they hold students to high expectations around effort, understanding, persistence, and performance in class.
- *Professional Learning about SEL*: The amount and quality of professional growth and learning opportunities available to faculty related to SEL.
- *Educating All Students*: Their readiness to address issues of diversity in the classroom.
- *Staff Engagement*: The supportiveness of their work environment.

Research questions

To inform DCPS's understanding of the role of school surveys in evaluating school leader's leadership and performance, this report addresses the research questions described below. The first question is focused on identifying whether staff surveys capture meaningful differences between schools. The second and third questions focus on understanding how much these differences might be due to principals, as opposed to other differences between schools. The fourth question examines the relationships between existing school leader evaluation scores and staff surveys.

1. How do staff survey scale scores vary within and between schools, and can they meaningfully differentiate schools from one another?

For the staff survey results to be informative for principal evaluations, they should be able to differentiate among schools when aggregated to the school level. A measure that does not differ between schools cannot distinguish between principals.

School-level constructs that are more immediately amenable to the influence of principals might be more easily differentiated between schools than constructs over which a principal's influence is less direct or immediate. Since principals may more immediately affect the school environment than teaching practices,

survey domains or scales that evaluate teachers' perceptions of the school environment may be more likely to differ at the school level than those evaluating teacher perceptions of students or of their own teaching practice. For example, the Insight survey measures teachers' perceptions of school instructional culture and may be more likely to differentiate schools than some of the Panorama scales, which measure teacher perceptions of student characteristics (such as the Perseverance scale) or teacher perceptions of their own practice (such as the Rigorous Expectations scale and the Educating All Students scale). The two Panorama scales measuring staff perspectives on the school environment (Professional Learning about SEL and Staff Engagement) may be more likely to differentiate schools than the other Panorama scales.

2. How stable are staff survey scores' year-to-year correlations for years in which the school's leader was the same versus years in which the school's leader was different?

Year-to-year correlations in school averages of staff survey domain and scale scores reflect how stable they are over time. To understand if staff survey scores may be capturing the effects of principals, it is useful to examine whether they are less stable over adjacent years when a principal changes compared to when the principal is the same. If they are more stable when the principal does not change, it suggests that these scores are in part driven by principals and that some of their stability over time is due to the effect of having the same principal.

3. Which survey scale scores respond more to changes in school leadership?

The third question further explores whether the differences in teacher surveys across schools might be due to principals, as opposed to other factors about the schools, and it identifies the survey scales or domains on which principals have larger effects. For this question, we examine what happens to staff survey measures when a principal enters or exits a school. If principals have effects on staff survey measures, we should observe changes in the staff survey measures when principals change schools. We also compare the changes to the staff survey measures to the changes in the percentage of students who are proficient in math and English language arts (ELA) to provide a benchmark to determine whether principals have bigger initial effects on staff surveys or student achievement.

4. What is the relationship between School Leader IMPACT components and relevant staff survey measures of the school environment?

The fourth question asks how the survey measures relate to existing principal evaluation components. For survey scales or domains that are conceptually related to the concepts that the principal evaluation system is aiming to assess (that they have face validity), a higher correlation between a survey measure and a principal evaluation component indicates that they are measuring related constructs, while a lower correlation indicates less related constructs. Survey measures could improve the overall validity of the evaluation system if they are correlated with (and conceptually related to) components of School Leader IMPACT (though very high correlations would mean the measures are not providing new information). In cases where survey measures have lower correlations with the School Leader IMPACT scores, DCPS should evaluate if the survey measures are measuring constructs that they believe should be but are not currently included in the principal evaluation system.

For each question, we focus on principals rather than assistant principals, because assistant principals may lack both the authority and the power to measurably affect the school environment as a whole, particularly if they are expected to focus on a specific aspect of school operations (which might differ at different schools).

The data sources, sample, and methods used in the analysis are in box 2.

Box 2. Data sources, sample, and methods

Data sources. The data for this study came from the District of Columbia Public Schools (DCPS) and the National Center of Education Statistics' Common Core of Data (CCD). There are five main types of data. First, we use principal IMPACT scores from 2015/16 through 2018/19 (including the final score), the Leadership Framework score, Student Outcome Goals score, and the scores for each of the six standards used to make the Leadership Framework score. Second, we have data from DCPS for the Panorama staff survey for 2017/18 and 2018/19, including responses from instructional staff to each item on the survey and information about the instructional staff that were eligible to take the survey, such as their years of employment and race/ethnicity. Third, we have data from DCPS for the Insight survey. This includes teacher responses to each item in 2017/18 and 2018/19 and school-level average scores for each domain and the index from 2015/16 to 2018/19. Although the Insight survey is offered in both fall and spring, we use the spring results for all analyses, as this is consistent with the timing of the Panorama survey. Fourth, we also use school data from the National Center of Education Statistics' CCD on school characteristics in each year from 2015/16 to 2018/19. Last, we have data from DCPS on student achievement on the Partnership for Assessment of Readiness for College and Careers test for each school from 2015/16 to 2018/19.

Sample. For research question 1 (How do staff survey scale scores vary within and between schools and can they meaningfully differentiate schools from one another?), we use teacher responses to the Insight and Panorama surveys for both 2017/18 and 2018/19 for the first part of the question and 2018/19 for the second part to test whether survey measures can differentiate schools within a given year. All teachers who responded to a given item are included in calculating the school-level averages of favorable responses for each item, and all teachers with a valid response to at least two items in a scale or domain are used in calculating school-level averages for each scale or domain.

For research question 2, we constrain the sample to those schools with responses for all Insight domains in three pairs of adjacent school years (see the top panel of table C4 in appendix C): 2015/16 and 2016/17, 2016/17 and 2017/18, and 2017/18 and 2018/19. (In the main report, we present findings only for the nine Insight domains for which we have four years of data. The remaining analyses for the other two Insight scales and for all the Panorama scales are in appendix C and are considered exploratory analyses, due to the small sample size.)

For research question 3, we use a principal-year dataset that includes all principals serving at schools that experienced any leadership transition during the four school years (2015/16 to 2018/19) and that have Insight survey data. This represents approximately 61 percent of principals in DCPS over this period of time. Because our study only has two years of data for two of the Insight domains and the Panorama scales, we focus this analysis on the nine Insight domains for which we have four years of data. A drawback of estimating only on a sample of principals who experienced a transition is that those who did and did not transition have different characteristics, and thus principal effects could differ for principals who never transition during the analysis period.

For research question 4, we use the sample of all schools with a principal who has a valid principal IMPACT score for the final score, Student Outcome Goals score, and Leadership Framework score and has a survey response for every Insight index or domains for 2015/16 to 2018/19 or Panorama scales for 2017/18 and 2018/19. See table B2 in appendix B for additional information.

Methodology. For research question 1, we calculate the intraclass correlations for each Panorama scale and Insight domain. An intraclass correlation is the fraction of the total variance of a measure that is due to variation between schools, as opposed to within schools. It indicates how much consensus there is among staff in the same school in how they respond to the survey. We also calculate the share of teachers within each school who respond favorably to each item in the Panorama and Insight surveys and report the 25th and 75th percentiles of this distribution for each item. We also calculate the average share of favorable responses within a school across all of the items in each Panorama scale and Insight domain and report the 25th and 75th percentiles of these distributions. We define meaningful differences between the 25th and 75th percentiles as being greater than 15 percentage points, moderate differences as 10 to 15 percentage points, and modest differences as less than 10 percentage points. For research question 2, we calculate pairwise correlations between Insight domains and Panorama scales (see appendix C) for adjacent years. For research question 3, we estimate the change in each school-average survey domain or scale when a principal changes schools, controlling for school characteristics that are constant over time. In the regressions used to estimate the principal effects, each school-level Insight domain score is standardized within each year relative to the average of the school-level domain scores. The approach compares principals relative to other principals who

have served in the same set of schools, where a set is defined as any school that has had at least one principal move to another school in the set from 2015/16 to 2018/19. One limitation of this approach is that if there are other factors changing at a school at the same time as a principal transition, the effects of these changes would be included in the principal effects. For research question 4, we calculate pairwise correlations between principal IMPACT components and teacher survey responses. We define small correlations as 0.1 to less than 0.3, moderate as 0.3 to less than 0.5, and large as 0.5 or higher. A more detailed description of the study data sources, sample, and methods is in appendix B.

Findings

This section presents findings to address the study’s four research questions.

Most of the variation in teacher survey measures occurs among teachers within the same school, rather than across schools

For both the Panorama and Insight teacher surveys, there was more variation in the survey measures among teachers in the same school than across schools across both years that we studied (2017/18 and 2018/19) (see table C1 in appendix C). The differences in teachers’ responses within the same school could arise because teachers have different opinions about the same conditions in the school or could be due to “true” variation in school conditions across grades, subjects, or students. For the Insight survey, the proportion of the total variation in each domain that arises due to differences in teacher responses *across* schools as opposed to among teachers in the same schools—called an intraclass correlation—was 0.10 to 0.22. This means 78 to 90 percent of the variance in these measures occurs among teachers within the same school, and the remaining 10 to 22 percent of the variance occurs across schools. These intraclass correlations are similar in magnitude to intraclass correlations of student math and reading tests, which range from 0.14 to 0.27 across grades K–12 (Hedges & Hedberg, 2007).

For Panorama, three scales (Perseverance, Professional Learning about SEL, and Staff Engagement) had intraclass correlations of 0.12 to 0.14, while two scales (Educating All Students and Rigorous Expectations) had intraclass correlations of 0.03 and 0.04. The low intraclass correlations for these two scales could be due to the fact that they are measuring teacher perceptions of their own practice (that is, whether they hold students to high expectations and whether they are prepared to address issues of diversity in the classroom), as opposed to the school environment.

Teacher survey responses differ meaningfully across schools

Although most of the variation in survey domain and scale scores occurs within schools, there are meaningful differences (greater than 15 percentage points) between the 25th and 75th percentiles of the school-level distribution in the share of teachers who have favorable responses to items on the Panorama and Insight surveys (figure 1).¹ On the Insight survey, there were large (at least 17 percentage points) differences between schools in the bottom quarter of the school-level distribution and the top quarter of the distribution in all domains (see figure 1; table C2 in appendix C). The differences were largest for the learning environment (32 percentage points) and leadership domains (24 percentage points).

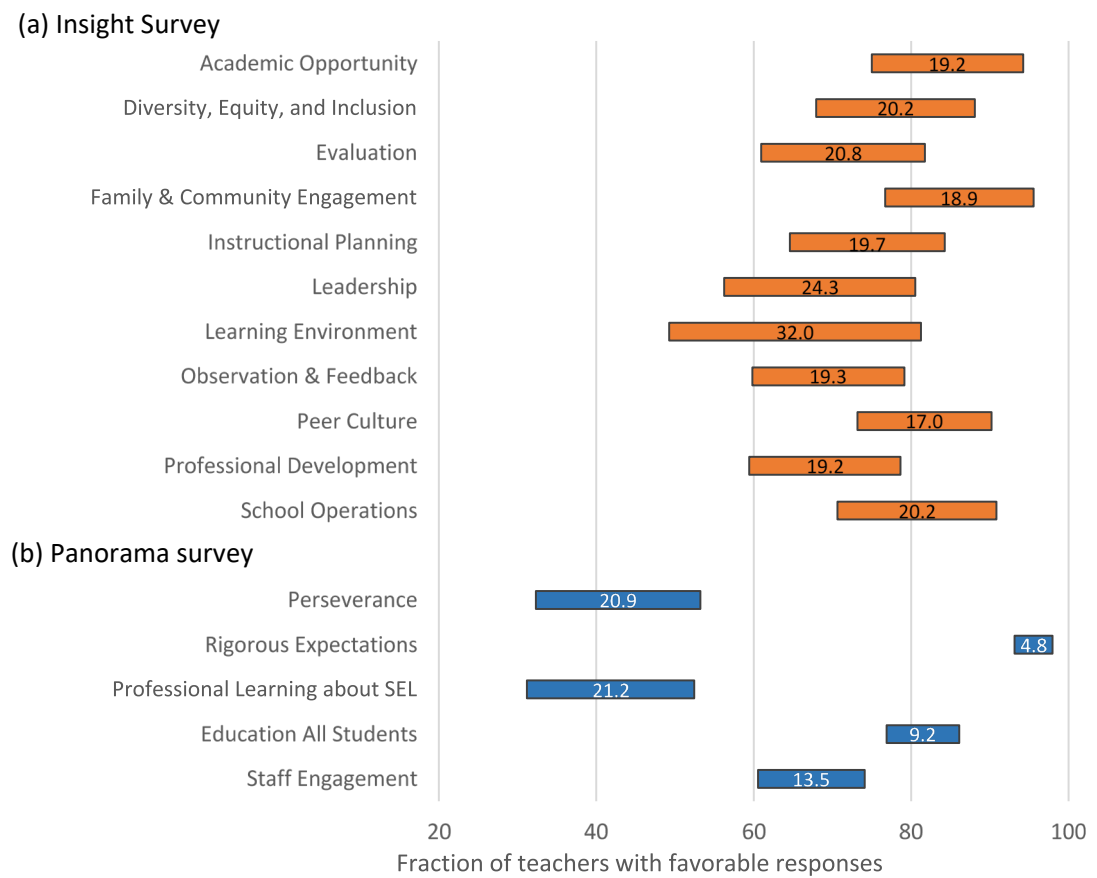
On the Panorama survey, there were large differences (greater than 20 percentage points) between schools in the bottom quarter of the distribution and the top quarter of the distribution for two of five scales: Perseverance and Professional Learning about SEL (see figure 1; table C3 in appendix C). For example, one item on the Professional Learning for SEL scale asks teachers: “In terms of social-emotional learning (SEL) in particular, how supportive has the school been of your growth as a teacher?” In one-fourth of schools, 36 percent or fewer

¹ In appendix C, we show all of the items on the Insight survey (table C2) and Panorama survey (table C3) DCPS offered in 2018/19, grouped by scale or domain.

teachers responded favorably to this item, while in another one-fourth of schools, 62 percent or more teachers responded favorably.

One of the remaining scales (Staff Engagement) had a moderate difference (14 percentage points) between schools in the bottom and top quarters of the distribution, while the other two had modest differences (less than 10 percentage points). Some items in the Staff Engagement scale had more variation across schools, while others had less. For example, in one-fourth of schools, less than 48 percent responded favorably to the statement, “My performance is measured against outcomes and metrics that are clearly explained,” while in another one-fourth of schools, at least 70 percent of teachers responded favorably (see table C3 in appendix C).

Figure 1. For all Insight domains and two Panorama domains (Perseverance and Professional Learning about SEL), interquartile ranges for the fraction of teachers in each school with favorable responses for each domain or scale were greater than 15 percentage points (2018/19)



SEL is social-emotional learning.

Note: The figure shows for each Insight domain and Panorama scale the interquartile range (25th to 75th percentiles) of the school distribution of the fraction of teachers in each school who respond favorably to each survey domain or scale (described in box 1).

Source: Authors' analyses based on survey data provided by the District of Columbia Public Schools, 2018/19.

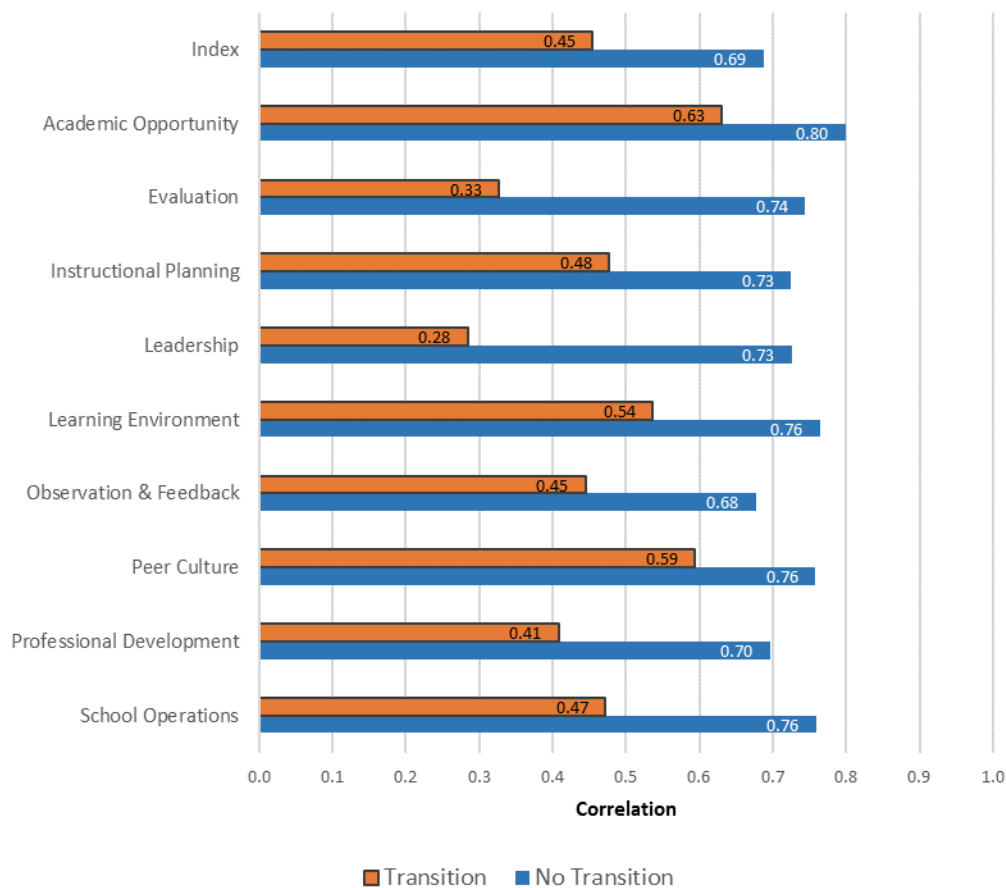
School-level averages of teacher Insight survey domains are consistently more stable across school years for schools with the same principal than across years when the principal changed

The average Insight survey responses of teachers within the same school are highly correlated across years in which a school leader did not change (figure 2). The correlations for all of the adjacent years in which a school leader did not change range from 0.68 to 0.80 for all survey measures. The year-to-year correlations are weaker on average for all Insight domains in the years in which a principal transitioned compared to when they did not.² The difference is largest for the Leadership domain, which we might expect would be most directly related to the leadership and performance of the principal. The pattern—of larger changes in domain scores when a school

² We report year-to-year correlations in table C4 in appendix C for both Insight and Panorama. The principal transitions that occur in 2015/16 to 2016/17 and in 2016/17 to 2017/18 follow this pattern clearly, but the transitions from 2017/18 to 2018/19 do not.

changes principals—suggests that the year-to-year correlations for the Insight domains in part may be measuring effects of principals.³

Figure 2. The year-to-year correlations for Insight domains are typically higher when schools do not experience a principal transition than when they do, particularly for Leadership and Evaluation (2015/16 to 2018/19)



Note: The figure shows the pairwise correlation coefficients between adjacent years for all adjacent year pairs (for the years noted in the legend) for each Insight domain score. We report results separately for adjacent years in which a school experienced a principal transition and for those in which a school did not. For Insight, the available years are 2015/16 to 2018/19. See table C4 in appendix C for details.

Source: Authors’ analyses based on survey data provided by the District of Columbia Public Schools, 2015/16, 2016/17, 2017/18, and 2018/19.

³ Because our study only has two years of data for two of the Insight domains and the Panorama scales, we focus on the results for the nine Insight domains for which we have four years of data for this analysis.

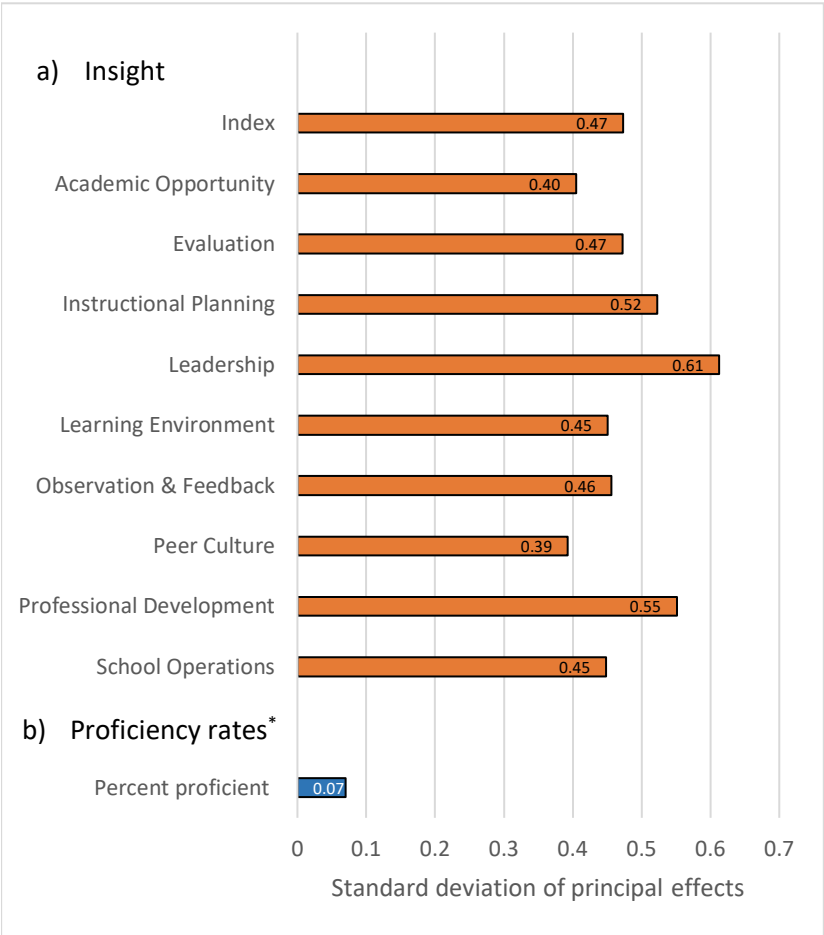
School Insight domain scores change substantially when principals change schools, with the largest changes in the Leadership domain. These changes are much larger than changes in student proficiency rates, suggesting that principals may have more immediate effects on constructs measured in teacher surveys—such as school climate—than on student achievement

To use teacher surveys for principal evaluation, it is important to understand whether principals affect teacher survey responses and in what survey domains or scales they have the largest effects.⁴ Controlling for school-level factors that are constant over time (such as the characteristics of the school neighborhood), we examine changes in teacher survey scores when principals switch schools—an approach that has been used to estimate principals’ effects on student test scores (Chiang et al., 2016; Grissom et al., 2015; Lipscomb et al., 2012).

When principals switch schools, school scores change most on the Leadership, Instructional Planning for Student Growth, and Professional Development domains. For all nine Insight domains and the index, the true standard deviations of the principal effects range from 0.39 to 0.61—while the standard deviation of principals’ effects on the percentage of students achieving proficiency on the Partnership for Assessment of Readiness for College and Careers (PARCC) assessments is only 0.07 (figure 3; see table C5 in appendix C).⁵ This implies that when principals switch schools, they have much larger immediate effects on staff survey measures than student proficiency rates.

To put these standard deviations in context, we can convert standard deviations to percentiles (see Appendix E for additional examples). Compared with the median principal in the district, a

Figure 3. The principal effects on the Insight domains are larger than the principal effects on proficiency rates in English language arts and math in the first years of a principal’s tenure (2015/16 to 2018/19)



*Percent proficient is the school average of the English language arts and math proficiency rates on the Partnership for Assessment of Readiness for College and Careers exam.

Note: The figure shows the true standard deviation of the principal effects on each Insight domain and the average percent of students who are proficient in English language arts and math. The true standard deviation is the variance of the fixed effects minus the mean of the squared standard errors. The principal effects are estimated with the set of schools who experienced a principal transition from 2015/16 to 2018/19.

Source: Authors’ analyses based on survey data provided by the District of Columbia Public Schools, 2015/16 to 2018/19.

⁴ It is also important that teacher survey responses provide valid indications of principal performance, which this study assesses only by examining their relationship to other measures of principal performance (that is, IMPACT scores).

⁵ Because our study only has two years of data for two of the Insight domains and the Panorama scales, we focus on the results for the nine Insight domains for which we have four years of data for this analysis.

principal at the 90th percentile of the distribution of principal effects on the Leadership domain would move the median school's Leadership domain score from the 50th percentile to the 78th percentile. In contrast, the standard deviation of the principal effects on percent proficient is only 0.07. A principal at the 90th percentile would move the median school's percent proficient rate from the 50th percentile to the 54th percentile.

All of the Insight domains and four Panorama scales have moderate or small correlations with the overall principal IMPACT score

Seven of the Insight domains and the index are moderately correlated with the total principal IMPACT score, while five have small correlations (see table C6 in appendix C). Three domains—Academic Opportunity, Learning Environment, and School Operations—have the highest correlations (0.37 to 0.39), while Observation & Feedback has the lowest correlation (0.20). These patterns are consistent for each year from 2015/16 to 2018/19, but the correlations are weaker for all domains and the Insight index in 2018/19.

One of the Panorama scales, Perseverance, is moderately correlated with the total principal IMPACT score (see table C6 in appendix C). Three scales have small correlations ranging from 0.19 to 0.28. The Educating All Students scale is not significantly correlated with the overall School Leader IMPACT score. These patterns are consistent for 2017/18 and 2018/19, the two years for which we had data for the Panorama survey and the principal evaluation scores.

All of the Insight domains and the three Panorama scales that have the highest correlations with the final School Leader IMPACT score are more correlated with the Leadership Framework scores than the Student Outcome Goals scores

The Leadership Framework score has larger correlations, on average, with the Insight index and domain scores than the Student Outcome Goals score, which is based on student achievement (see table C7 in appendix C). Across all the Insight domains and the index, the correlations with the Student Outcome Goals scores are small, ranging from 0.09 to 0.21, while the correlations with the Leadership Framework scores are mostly moderate, ranging from 0.26 to 0.46.⁶ This pattern might be expected because the preceding evidence on principal transitions suggests that principals have larger effects on Insight teacher survey scores than on student proficiency. This finding suggests that the Insight survey is capturing complementary information to the Leadership Framework score, with a different perspective on principals' leadership and performance. We are aware of no other studies that have examined the relationship between teacher surveys and ratings of principals' professional practice. However, there have been analogous studies correlating *student* surveys with ratings of *teachers'* professional practice, which have found correlations that are similar or slightly smaller than those found here (Chaplin et al., 2014).

Similar to the Insight survey, the three Panorama scales with the strongest correlations with the total principal IMPACT score are also more correlated with the Leadership Framework scores than with the Student Outcome Goals scores (see table C7 in appendix C). For the Perseverance, Professional Learning about SEL, and Staff Engagement scales, the correlations between the Panorama scales and the Leadership Framework scores are small or moderate, ranging from 0.27 to 0.38. The other two scales—Rigorous Expectations and Educating All

⁶ As discussed in box 1, the total principal IMPACT score is composed of two component scores: the Leadership Framework score and the Student Outcome Goals score. The Leadership Framework score evaluates school leaders on six standards (Instruction, Talent, School Culture, Operations, Family & Community, and Personal Leadership), while the Student Outcome Goals score measures student learning outcomes, such as proficiency rates on standardized tests. The Leadership Framework score consists of two cycles: one that occurs midyear (cycle 1) and another at the end of the year (cycle 2). The correlations between each Leadership Framework cycle score and the Panorama survey scales and Insight domains are very similar, though they are slightly higher on average for the cycle 2 score. This may be because the Panorama survey and the Insight survey we use in this study are offered in the spring, which is closer to the cycle 2 evaluation.

Students—have very small, insignificant correlations (less than 0.1) with the Leadership Framework. The three Panorama scales that have small or moderate correlations with the Leadership Framework scores are also the Panorama scales that are most correlated with the Insight scores (see table C9).

The Insight domains and the three Panorama scales that are correlated with the Leadership Framework scores have larger correlations, on average, with three standards within the Leadership Framework: Instruction, Talent, and School Culture

For the scores on each of the six standards within the Leadership Framework score, the Insight domains and the index have larger correlations on average with the Instruction, School Culture, and Talent standards than the other standards (see table C8 in appendix C). Over 75 percent of the correlations between the Insight domains and these three standards are over 0.2, and one-third of the correlations are over 0.3. The Insight domains' correlations with the other standards—Operations, Family & Community, and Personal Leadership—rarely exceed 0.2.

The three Panorama scale scores that are correlated with the Leadership Framework scores also have larger correlations with the Instruction, Talent, and School Culture scores within the Leadership Framework, on average (see table C8 in appendix C). The correlations between the Perseverance, Professional Learning about SEL, and Staff Engagement scales and the Instruction, Talent, and School Culture standards within the Leadership Framework range from 0.24 to 0.52.

Limitations

This study has several limitations. First, it is important to caution that this study uses teacher survey data collected during a time that the surveys are not being used directly for principal evaluation. If the teacher surveys start to be used for principal evaluation, it is possible that the survey responses could change in ways that undermine their usefulness.

Second, the Panorama survey and two of the Insight domains are only available for two years. We do not present results for these scales for research questions 2 or 3 due to the small number of principal transitions that occurred over this period of time.

Third, one limitation for research question 3 is that the principal effects can only be estimated for the principals that switched schools at any point from 2015/16 to 2018/19; this constitutes approximately 61 percent of the principals who served in DCPS schools from 2015/16 to 2018/19 (see table B3 in appendix B). Principals who changed schools tend to have different characteristics than those who do not. For example, they have lower average School Leader IMPACT evaluation scores and serve schools with a higher fraction of students eligible for free lunch. Given these differences, it is possible that principal effects could differ for principals who never transition schools from 2015/16 to 2018/19. A fourth limitation, also related to research question 3, is that if there are other factors changing at a school at the same time as a principal transition, the effects of these changes would be included in the principal effect, even though they may not be caused by the principal.

Finally, supervisors evaluating principals on the Leadership Framework standards might be influenced by previous teacher survey results. Supervisors do not typically have access to the results of the Insight fall survey when conducting the midyear assessment or the spring survey when conducting the end-of-year assessments. However, supervisors may look at the data from prior surveys when assessing principals. While the survey results are not intended to directly inform a principal's evaluation, supervisors may include principals' responses or actions taken, based on the data when making their assessments. If so, this could artificially inflate the correlation between the Leadership Framework scores and teacher surveys.

Implications

This study describes the properties of staff surveys for the purposes of principal evaluation. Taken together, the findings suggest that staff surveys and the teacher perspectives they represent can offer new and potentially actionable information to the principal evaluation process.

School averages of teachers' responses on all of the Insight domains and three of the Panorama scales differ substantially between schools at the 25th and 75th percentiles. Our findings suggest that staff surveys that are designed to capture features of the school environment, such as the Insight survey and two of the Panorama scales that differ across schools (Staff Engagement and Professional Learning for SEL), provide information that can differentiate principals and could be used to inform principal feedback by identifying strengths or areas for improvement.

The fact that year-to-year correlations in staff survey scores are weaker for adjacent years in which a school gets a new principal suggests that part of what contributes to year-to-year stability in correlations is having the same principal. In other words, it suggests that principals affect these measures. We observe this pattern for Insight, but small sample sizes impeded our ability to robustly assess this relationship with Panorama and two of the Insight domains that were only offered for two years. In future work, it would be useful to collect additional years of survey data to examine if the pattern holds in these other scales.

Another way to understand whether principals affect staff surveys is to examine what happens to the staff survey scores when principals change schools. We find that when principals change schools, school-level averages of the Insight teacher survey domains change more than the percentage of students scoring proficient on the PARCC exam. This finding suggests that principals have the capacity to change important aspects of the school environment, even in their first year in a school—sooner than they could produce effects that are evident in student achievement.

Our finding that the Insight domains and three Panorama scale scores—Perseverance, Professional Learning about SEL, and Staff Engagement—are correlated with the overall principal IMPACT scores (particularly with the Leadership Framework half of the principal IMPACT score) suggests that they are measuring something that is related to what is measured in the Leadership Framework. However, the correlations never exceed 0.55, suggesting they are related to the existing School Leader IMPACT scores but not redundant. To the extent that the staff surveys are indeed accurately reflecting principal quality, then including these measures in the principal evaluation system would therefore add information to the evaluation system.

Taken together, the nine Insight survey domains (for which we have four years of data) have properties indicating they differentiate between schools, are affected by principals, and are related to, but not redundant with, the existing evaluation system. To the extent that school leaders believe the constructs measured by these domains are measuring features of the school environment that they would like principals to affect, these measures could be a useful addition to the principal evaluation system. Three Panorama scales which differentiate schools and are correlated with principal IMPACT scores (Perseverance, Professional Learning about SEL, and Staff Engagement) and the two Insight domains offered for only two years in our data (Diversity, Equity, & Inclusion and Family & Community Engagement) could potentially be useful to include in the principal evaluation system. However, we were not able to include these scales and domains in our analyses for research questions 2 and 3 and would need to do so (requiring additional years of data) before recommending their use.

References

- Burkhauser, S. (2017). How much do school principals matter when it comes to teacher working conditions? *Educational Evaluation and Policy Analysis*, 39(1), 126–145.
- Chaplin, D., Gill, B., Thompson, A., & Miller, H. (2014). *Professional practice, student surveys, and value-added: Multiple measures of teacher effectiveness in the Pittsburgh Public Schools* (REL 2014–024). Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. <https://eric.ed.gov/?id=ED545232>
- Chiang, H., Lipscomb, S., & Gill, B. (2016). Is school value added indicative of principal quality? *Education Finance and Policy*, 16(3), 283–309.
- Coelli, M., & Green, D. (2012). Leadership effects: School principals and student outcomes. *Economics of Education Review*, 31(1), 92–109.
- Dhuey, E., & Smith, J. (2014). How important are school principals in the production of student achievement? *Canadian Journal of Economics*, 47(2), 634–663.
- Garet, M. S., Wayne, A. J., Brown, S., Rickles, J., Song, M., Manzeske, D., & Ali, M. (2017). *The impact of providing performance feedback to teachers and principals*. Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. <https://ies.ed.gov/ncee/pubs/20184001/>
- Grissom, J., Kalogrides, D., & Loeb, S. (2015). Using student test scores to measure principal performance. *Educational Evaluation and Policy Analysis*, 37(1), 3–28. <https://doi.org/10.3102/0162373714523831>
- Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis*, 29(1), 60–87.
- Leithwood, K., Seashore Lewis, K., Anderson, S., & Wahlstrom, K. (2004). *Review of research: How leadership influences student learning*. The Wallace Foundation. <https://www.wallacefoundation.org/knowledge-center/Documents/How-Leadership-Influences-Student-Learning.pdf>
- Lipscomb, S., Chiang, H., & Gill, B. (2012). *Value-added estimates for phase 1 of the Pennsylvania teacher and principal evaluation pilot: Full report*. Mathematica Policy Research. <https://eric.ed.gov/?id=ED531795>
- Liu, K., Stuit, D., Springer, J., Lindsay, J., & Wan, Y. (2014). *The utility of teacher and student surveys in principal evaluations: An empirical investigation*. (REL 2015–047). Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. <https://eric.ed.gov/?id=ED548024>
- Sebastian, J., & Allensworth, E. (2012). The influence of principal leadership on classroom instruction and student learning: A study of mediated pathways to learning. *Educational Administration Quarterly*, 48(4), 626–663.

August 2021

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-17-C-0006 by Regional Educational Laboratory Mid-Atlantic administered by Mathematica. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

This REL report is in the public domain. While permission to reprint this publication is not necessary, it should be cited as:

Kozakowski, W., Gill, B., & Shiferaw, M. (2021). *Exploring the potential role of staff surveys in school leader evaluation* (REL 2021-117). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. <http://ies.ed.gov/ncee/edlabs>

This report is available on the Regional Educational Laboratory website at <http://ies.ed.gov/ncee/edlabs>.

Exploring the Potential Role of Staff Surveys in School Leader Evaluation

Appendix A. About the study

Appendix B. Methods

Appendix C. Supporting analysis

Appendix D. Nonresponse bias analysis

Appendix E. Principal effect interpretations

Appendix A. About the study

This study builds from work on the School Leader IMPACT coaching project, a prior just-in-time project that the Mid-Atlantic Regional Educational Laboratory (REL) conducted with the District of Columbia Public Schools (DCPS). Findings from this prior project are shared in this appendix.

In spring 2019, REL Mid-Atlantic worked with DCPS to understand possible ways to enhance the measures in the School Leader IMPACT system to make them more reliable and valid (Shiferaw & Gill, 2019). We examined year-to-year changes in School Leader IMPACT scores and sub-scores to better understand the role of random variation in the rating system. In August 2019, we conducted a coaching session with DCPS to discuss the following core lessons from this exercise:

- Overall, School Leader IMPACT scores have been stable over the last several years. Variation is driven largely by scores on Student Outcome Goals rather than scores in the Leadership Framework component.
- In most years, there has been little systematic variation by grade level in Student Outcome Goals results. The Partnership for Assessment of Readiness for College and Careers (PARCC) scales and cut scores are consistent across grades.
- Annual changes in proficiency rates for the PARCC are not stable from year to year. Few school-level proficiency changes are statistically significant.
- The likelihood of erroneously identifying a school as having declining performance is substantial.

Based on the findings, we made the following recommendations:

- Incorporate additional measures of principal performance, based on perceptions of school staff. This will provide information on principals' professional practice from the perspective of those who see the principal every day and has the potential to provide useful feedback to improve performance.
- Average proficiency scores across multiple years to reduce random error in Student Outcome Goals scores.
- Use a value-added measure instead of change in proficiency rates.
- Incorporate Bayesian shrinkage to reduce random error in Student Outcome Goals scores.

Appendix B. Methods

This appendix provides a detailed description of the analyses used for this study. This description includes the data sources, variables, samples, and research methods, as well as a discussion of how missing data and survey nonresponse were addressed.

Data

The key data sources for this study are the District of Columbia Public Schools' (DCPS) The New Teacher Project's (TNTP) Insight teacher survey, Panorama Education (Panorama) staff survey, and principal evaluation records. It also includes school-level data on student achievement and school type from DCPS and on student demographics and enrollment from the National Center for Education Statistics' Common Core of Data (CCD).

Insight staff surveys. DCPS began working with the Instructional Culture Insight survey in 2015/16. The survey is a validated measure of school instructional culture created and administered by TNTP. DCPS administers the Insight survey in the fall and spring terms to all teachers in the district; we use the spring term for all analyses to be consistent with Panorama, which was also offered in the spring. The survey contains the following 11 domains and an index score (those marked with * refer to domains that were added in 2017/18):

- *Academic Opportunity:* Teachers' perceptions of whether rigorous academic practices are visible in classrooms and whether school leaders have the necessary resources and knowledge to support these practices.
- *Diversity, Equality, and Inclusion*:* Teachers' perceptions of whether diversity and inclusion are a clear priority at their school and whether the school's talent management practices and overall culture advance that priority.
- *Evaluation:* Teachers' perceptions of how clearly performance expectations have been communicated, how much teachers agree with those expectations, and how accurately evaluations represent teacher performance in this school.
- *Family & Community Engagement*:* Teachers' perceptions of the relationships they have built with students' families and the systems in place for keeping families involved in students' education.
- *Instructional Planning for Student Growth:* Teachers' perceptions of whether they share student assessments or receive instructional planning support to make sense of student assessment data and use it to improve their instruction.
- *Leadership:* Teachers' perceptions of the effectiveness of school leaders, including the vision they set for the school, the extent to which they communicate and follow through on clear goals and priorities, and whether they seek upward feedback from their teachers.
- *Learning Environment:* Teachers' perceptions of the consistency of expectations and consequences for student conduct, as well as leadership support in maintaining a productive learning environment.
- *Observation and Feedback:* Teachers' perceptions of how frequently they are observed and how consistently they get helpful feedback to improve their instruction.
- *Peer Culture:* Teachers' perceptions of colleagues within the school, including whether peers share similar priorities and provide examples of exemplary teaching.
- *Professional Development:* Teachers' perceptions of experiences with professional development and how well it has helped them learn new, relevant skills and improve their instructional practice.

- *School Operations*: Teachers' perceptions of the efficiency and effectiveness of daily operations in the school building, including non-academic services, student information systems, and overall cleanliness.
- *Index score*: A summary measure based on teacher agreement with three indicators of instructional culture: that the expectations of effective teaching are clearly defined, that there is a common vision of what effective teaching looks like, and that their school is committed to improving instructional practice.

For Insight, survey domain scores and the index score are calculated by the survey developer (TNTP) and shared with DCPS. Response options are scales measuring levels of agreement with each item, with answers of 0 indicating strong disagreement and 5 indicating strong agreement.

We received school-level Insight domain scores and the index score for 2015/16 through 2018/19 and teacher-level responses to the Insight survey in 2017/18 and 2018/19.

We also received data from DCPS on the set of teachers eligible to take the Insight survey in 2017/18 and 2018/19.

Panorama staff surveys. DCPS began working with Panorama Education in 2017/18 to administer the Panorama survey each spring to students, staff, and parents to measure their perspectives on social-emotional learning (SEL), school climate, and engagement. The staff survey consists of one set of items given to instructional staff (such as teachers, aides, and coaches) and another set of items for noninstructional staff. This study uses the instructional staff survey.

The instructional staff survey (which we refer to as the teacher survey) consists of 43 questions. Each question is part of one of the following five scales (scales with an asterisk were created for or modified by DCPS):

- *Perseverance*: Teachers' perceptions of how well their students can persevere through setbacks to achieve important long-term goals.
- *Rigorous Expectations**: Teachers' perceptions of how much they hold students to high expectations around effort, understanding, persistence, and performance in class.
- *Professional Learning about SEL*: Teachers' perceptions of the amount and quality of professional growth and learning opportunities available to faculty related to SEL.
- *Educating All Students**: Teachers' perceptions of their readiness to address issues of diversity in the classroom.
- *Staff Engagement**: Teachers' and staff's perceptions of the supportiveness of the work environment.

The teacher survey scales measure teachers' perceptions of students' SEL competencies (such as Perseverance) and their own competencies (such as holding students to Rigorous Expectations and Educating All Students). Two of the scales measure features of the school environment, including Staff Engagement and opportunities for Professional Learning about SEL. Three of the scales—Rigorous Expectations, Educating All Students, and Staff Engagement—were created specifically for DCPS.

For each question, the responses are values ranging from 1 to 5, typically indicating levels of agreement or frequency, with answers of 1 indicating less agreement or frequency and 5 indicating more agreement or frequency. The study team created scale scores by averaging a teacher's responses across all of the items in the scale, so the resulting scales range from 1 to 5 points. Following DCPS's methodology, this study counted the scale as non-missing if a respondent completed at least two items in the scale.

We use Panorama teacher survey data for two school years: 2017/18 and 2018/19. There were 3,106 teachers in 2017/18 and 3,121 teachers in 2018/19 who responded to at least one item on the survey.

We also received data from DCPS on the set of instructional staff eligible to take Panorama in 2017/18 and 2018/19. These data included the gender, years of experience, and race/ethnicity of the staff member. Instructional staff include teachers as well as coaches, aides, and residency fellows.

Principal evaluations. The DCPS evaluation system (called School Leader IMPACT) measures school leaders' effectiveness with the goals of setting clear expectations and providing feedback for improvement. The principal evaluation data we use includes a school leader's final score (a summative measure of principal performance), the final rating a principal receives (minimally effective, effective, or highly effective), and subcomponent scores for all principals and assistant principals from 2015/16 to 2018/19. The data elements are listed in table B1. Unless otherwise noted, all analyses use the principal evaluations scores, not the assistant principal scores.

The final score is an equally weighted average of two components scores: the Leadership Framework score and the Student Outcome Goals score.

The Leadership Framework aims to assess the effectiveness of a school leader's practices with respect to improving student learning, as gauged by the school leader's supervisor. The Leadership Framework assesses school leaders relative to six standards (described below), using both qualitative and quantitative measures. The six standards measure the extent to which the school leader does the following:

- *Instruction:* Develops the school's instructional vision and goals, oversees effective school- and classroom-level planning, ensures effective classroom instruction, and establishes a culture of data-driven instruction.
- *Talent:* Identifies and strategically places outstanding talent, evaluates staff members, provides support, removes low performers, retains key staff, and builds leadership capacity.
- *School Culture:* Creates a positive, student-centered environment; ensures students meet high academic and behavioral expectations; and implements effective interventions that support student success.
- *Operations:* Efficiently manages school operations, maximizes impact of limited resources, and fulfills all legal and policy requirements.
- *Family & Community:* Builds relationships with families and community members, efficiently responds to families' inquiries and concerns, and shares information with families to support their children's success.
- *Personal Leadership:* Engages in continuous self-improvement, communicates effectively, demonstrates cultural competence, and perseveres in the face of obstacles.

There are two assessment cycles used to calculate the Leadership Framework score: a midyear evaluation, which makes up 20 percent of the final IMPACT score, and end-of-year evaluation, which makes up 30 percent of the final IMPACT score. In each evaluation, principals receive a rating from their supervisors (ranging from 1 to 4) for each standard. For the analysis, we use the weighted Leadership Framework score and an average score for each standard in which we average the midyear and end-of-year assessments.

The Student Outcome Goals score measures school-level learning outcomes for students. There are two components to the Student Outcome Goals measures: student achievement goals and school-specific goals. The student achievement goals are based on improving student proficiency and reducing the percentage of students scoring at the lowest levels of the standardized test used for accountability purposes (the Partnership for Assessment of Readiness for College and Careers, or PARCC). If PARCC is unavailable for a school, other student achievement measures are used. The student achievement goals constitute 30 percent of the final IMPACT score. The school-specific goals are goals that school leaders set that address high-need areas for their school's overall success. They can measure student achievement or focus on a part of school culture that will help support student

learning. School-specific goals make up 20 percent of the final IMPACT score. For the analysis, we use the summative Student Outcome Goals score, not the component scores.

School data. The school data come from two sources: DCPS and the National Center for Education Statistics' CCD. The data elements are listed in table B1. DCPS provided data for each school and year from 2015/16 to 2016/17 on the percentage of students who are proficient in math and English language arts on the district's standardized test (PARCC) at each school for each year. They also provided information for each school about the type of school (elementary, education campus, middle, high, or other). The study team also used data from CCD for each school in each year from 2015/16 to 2018/19 on enrollment; the share of students who were Black, Hispanic, and White; and the share who directly certified as eligible for the federal school lunch program.

Table B1. Data elements used for principal evaluation and school characteristics (2015/16 to 2018/19)

Characteristic	Description
Principal evaluation variables	
Final score	The summative score for each principal for the principal IMPACT evaluation system.
Final rating	The final rating for each principal, based on the final score. It has the following values: <ul style="list-style-type: none"> Minimally effective Effective Highly effective
Student Outcome Goals score	The weighted score for whether the principal has met his or her student goals, which include student achievement goals (such as PARCC exam proficiency rates) and school-specific goals.
Leadership Framework score	The weighted average of the scores on each of the six Leadership Framework standards from the midyear and end-of-year evaluation cycles. The midyear score represents 40 percent of the score and the end-of-year score represents the other 60 percent of the score.
Instruction score	The average score on the Instruction standard of the Leadership Framework from the midyear and end-of-year evaluations.
Talent score	The average score on the Talent standard of the Leadership Framework from the midyear and end-of-year evaluations.
School Culture score	The average score on the School Culture standard of the Leadership Framework from the midyear and end-of-year evaluations.
Operations score	The average score on the Operations standard of the Leadership Framework from the midyear and end-of-year evaluations.
Family & Community score	The average score on the Family & Community standard of the Leadership Framework from the midyear and end-of-year evaluations.
Personal Leadership score	The average score on the Personal Leadership standard of the Leadership Framework from the midyear and end-of-year evaluations.
School data	
Percent proficient math	Percentage of students in a school who score at level 4 or 5 on the PARCC math exam.
Percent proficient ELA	Percentage of students in a school who score at level 4 or 5 on the PARCC ELA exam.
Percent direct certified	Percentage of students in a school who directly certify for free lunch.
Total enrollment	Total number of students enrolled in a school.
Percent Black	Percentage of students in a school who are Black.
Percent Hispanic	Percentage of students in a school who are Hispanic.
Elementary school	Indicator for whether a school is an elementary school.
Education campus	Indicator for whether a school is an education campus.
Middle school	Indicator for whether a school is a middle school.
High school	Indicator for whether a school is a high school.
Other program	Indicator for whether a school is any other kind of school, such as an alternative school.

ELA is English language arts. PARCC is Partnership for Assessment of Readiness for College and Careers.

Sample

The sample sizes differed depending on the research question and the analysis. We describe the samples used for each research question and analysis in table B2.

Table B2. Description of samples used for each research question and analysis

Research question (RQ)	Table/analysis	Analysis sample	Sample size
RQ1. How do staff survey scale scores vary within and between schools and can they meaningfully differentiate schools from one another?	Intraclass correlations (2017/18, 2018/19)	Teachers eligible for Panorama survey or Insight survey who completed it in 2017/18 and 2018/19	Samples below vary across items: 2,712 to 2,866 teachers for Insight, 2017/18 2,584 to 2,671 teachers for Insight, 2018/19 3,085 to 3,100 teachers for Panorama, 2017/18 3,123 to 3,126 teachers for Panorama, 2018/19
	School-level rates of favorable responses on the Panorama and Insight surveys (2018/19)	Schools eligible for Panorama survey or Insight survey who completed it in 2018/19	115 schools for Insight 117 schools for Panorama
RQ2. How stable are staff survey scores' year-to-year correlations for years in which the school's leader was the same versus different?	Year-to-year correlations for adjacent years in which schools did or did not experience a principal transition	Schools eligible for Panorama survey or Insight survey who completed it in 2015/16, 2016/17, 2017/18, and 2018/19	1) <i>No principal transition:</i> Insight: 83 schools for 2015/16 to 2016/17 89 schools for 2016/17 to 2017/18 88 schools for 2017/18 to 2018/19 260 schools for 2015/16 to 2018/19 Panorama: 91 schools for 2017/18 to 2018/19
			2) <i>Principal transition:</i> Insight: 18 schools for 2015/16 to 2016/17 19 schools for 2016/17 to 2017/18 20 schools for 2017/18 to 2018/19 57 schools for 2015/16 to 2018/19 Panorama: 20 schools for 2017/18 to 2018/19
RQ3. Which survey scale scores respond more to changes in school leadership?	Principal effects on teacher surveys	For 9 Insight domains offered from 2015/16 to 2018/19: Any schools that experienced a principal change from 2015/16 to 2018/19	195 principal-years
RQ4. What is the relationship between School Leader IMPACT components and relevant staff survey measures of the school environment?	Correlations between staff surveys and principals' total IMPACT score (2015/16 to 2018/19) and correlations between staff surveys and principals' IMPACT subscores (2015/16 to 2018/19)	All principals from 2015/16 to 2018/19 with a final evaluation score, Student Outcome Goals score, and Leadership Framework score who completed the Insight or Panorama survey	For Insight, 9 domains: 418 principal-years for 2015/16 to 2018/19 For Insight, 2 domains: 213 principal-years for 2017/18 to 2018/19 For Panorama: 215 principal-years for 2017/18 to 2018/19
	Correlations between staff surveys and principals' Leadership Framework subscores (2018/19)	All principals with a Leadership Framework score in 2018/19 who completed the Insight or Panorama survey	For Insight: 106 principals For Panorama: 108 principals

Analysis methods

Research question 1. How do staff survey scale scores vary within and between schools and can they meaningfully differentiate schools from one another?

The study team calculated intraclass correlation coefficients using teacher-level survey responses to the Insight and Panorama surveys for the two years in which teacher-level data was available (2017/18 and 2018/19). For each item and scale or domain on the Panorama and Insight surveys, the team also calculated the percentage of teachers with favorable responses in each school in 2018/19. A favorable response was a 4 or 5 on the Panorama survey (which has a scale from 1 to 5) and a 3, 4, or 5 on the Insight survey (which has a scale from 0 to 5). For each scale, we took the average rate of favorable responses at the school level across all items in the scale. We calculated the 25th and 75th percentiles of the distribution of the percentage of favorable responses in each school for each item and scale in 2018/19 (see tables C2 and C3 in appendix C). We define meaningful differences between the 25th and 75th percentiles as being greater than 15 percentage points, moderate differences as 10 to 15 percentage points, and modest differences as less than 10 percentage points.

Research question 2. How stable are staff survey scores' year-to-year correlations for years in which the school's leader was the same?

The study team calculated pairwise year-to-year correlations between adjacent years for school averages of teacher survey measures separately for the set of schools that experienced a principal transition over the adjacent years and for the set of schools that did not. For Insight, there were three adjacent-year pairs for nine of the domains and the index: 2015/16 and 2016/17, 2016/17 and 2017/18, and 2017/18 and 2018/19. For Panorama and two of the Insight domains (Diversity, Equity, and Inclusion; Family & Community Engagement), there was only one adjacent-year pair: 2017/18 and 2018/19. For Insight, we also calculated the year-to-year correlation using all adjacent-year pairs to get an average year-to-year correlation.

Comparing the year-to-year correlations when principals transitioned and when they did not provide insight into whether principals are part of what makes the year-to-year correlations more or less stable. If the correlations are less stable when a school leader has transitioned, it suggests that some of the stability of the year-to-year correlations for adjacent years in which principals did not change is coming from the persistent effect of the principal. Because the sample size is so small for Panorama and the two Insight domains (only 20 for years with a principal transition), it will be challenging to get an accurate estimate of the year-to-year correlations in the years with a principal transition. As a result, we focus on the results for Insight (excluding the two domains only offered for two years) in the main body of the report and keep these exploratory analyses in the appendix.⁷

Research question 3. Which survey scale scores respond more to changes in school leadership?

For this question, the study team estimated a principal value-added model that evaluates the within-school change in survey measures when there is a change in school leadership (Chiang et al., 2016; Grissom et al., 2015; Lipscomb et al., 2012). A school's learning environment could be affected by its principal or by other school-specific factors that are beyond the principal's control, such as neighborhood quality. When a principal remains in

⁷ Schonbrodt and Perugini (2013) provide guidance about the sample size needed for a correlation to stabilize across repeated draws from random samples. In this case, we will assume the true year-to-year correlation is about 0.5 in a year with a principal transition (which is the average correlation we observe for Insight for years with a principal transition) and 0.7 in a year without a principal transition (which is the average correlation we observe for Insight for years without a transition). Schonbrodt and Perugini's simulations suggest that a minimum sample size of 50 schools would be necessary for 90 percent of random samples to have a correlation that falls within the range of 0.3 to 0.7—the range that would be necessary for us to detect a difference between the years without a transition and those with one. Based on this guidance, we do not include analyses which use sample sizes of less than 50 in the main body of the report.

the same school, it is not possible to separate the school effect from the principal effect. However, when principals change schools, they potentially reveal the principal effect in the difference between outcomes for the same school with different principals. Using approximately four years of staff survey data, we observe what happens to survey outcomes in schools when there is a principal transition.

For this analysis, we use the Insight survey data which is available from 2015/16 to 2018/19. The Panorama survey and two Insight domains (Family & Community Engagement; Diversity, Equity, and Inclusion) are only available for two years. We do not include them in the analysis because the principal value-added estimates would be very noisy (that is, have random irregularity) and limited to the set of principals who transitioned schools between 2017/18 and 2018/19.

One important limitation of the principal value-added analysis is that we can only estimate principal effects for the set of leaders who transition schools from 2015/16 to 2018/19. We show the number of principals who transitioned schools in adjacent years over this period of time in table B3. For each year, 29 percent of principals transitioned schools, and 16 to 18 percent of schools experienced a principal transition. Across all years from 2015/16 to 2018/19, 61 percent of principals experienced a transition and 42 percent of schools experienced a principal change.

The principals who transitioned had lower average evaluation scores than those that did not transition (291 versus 298). In addition, a lower share of principals who transitioned were ever rated highly effective (table B4). Schools that experienced a leader change also had lower proficiency rates in math (19 percent versus 25 percent) and ELA (22 percent versus 28 percent) (table B5). These schools also had a higher share of students who were Black (75 percent versus 66 percent) or who directly certified for free and reduced-price lunch (53 percent versus 47 percent) and a lower share of Hispanic students (13 percent versus 17 percent) (see table B5). Also, schools that experienced a transition were less likely to be elementary schools (50 percent versus 58 percent) and were more likely to be high schools (21 percent versus 16 percent).

Table B3. Principal transitions from 2015/16 to 2018/19

Years	Principals			Schools		
	Total	Number who transitioned schools	Percentage who transitioned schools	Total	Number with a leader change	Percentage with a leader change
2015/16 and 2016/17	123	35	28.5	110	18	16.4
2016/17 and 2017/18	125	36	28.8	113	19	16.8
2017/18 and 2018/19	126	37	29.4	114	20	17.5
All years	160	97	60.6	114	48	42.1

Note: This table shows the total number of principals who transitioned schools during the noted years from 2015/16 to 2018/19 and the total number of schools that experienced a principal change during the years noted from 2015/16 to 2018/19. Schools are only counted as experiencing a transition if they are present in the data in both years, the principal changes, and the school has Insight survey data in both years.

Source: Authors' analyses based on administrative data and survey data provided by the District of Columbia Public Schools, 2015/16 to 2018/19.

Table B4. Characteristics of principals who transitioned schools (2015/16 to 2018/19)

Principal characteristics	All principals	Principals who transitioned schools from 2015/16 to 2018/19
Average total evaluation score	297.6	291.2
Share ever rated minimally effective	47.8	50.0
Share ever rated highly effective	34.0	22.9

Note: The table shows characteristics for the principals who transitioned schools during two time periods (2015/16 to 2018/19 and 2017/18 to 2018/19) relative to all principals in the district. Average total evaluation score is the average final score from the school leader evaluation system. Share ever rated minimally effective (or highly effective) is the fraction of principals that were ever rated as minimally effective (or highly effective) over the period of time noted in the column heading. Schools are only counted as experiencing a transition if they are present in the data in both years, the principal changes, and the school has Insight survey data in both years.

Source: Authors' analyses based on administrative data and survey data provided by the District of Columbia Public Schools, 2015/16 to 2018/19.

Table B5. Characteristics of schools with leadership transitions (2015/16 to 2018/19)

School characteristics	Total	Had a leader change from 2015/16 to 2018/19
Achievement		
Percent proficient in math	25.3	19.2
Percent proficient in ELA	27.6	21.5
Demographics		
Percent direct certified	47.2	52.6
Percent Black	66.3	74.7
Percent Hispanic	16.6	13.3
School type		
Percent elementary school	58.4	50.0
Percent education campus	13.3	12.5
Percent middle school	12.4	16.7
Percent high school	15.9	20.8
Percent other program	4.4	2.1

ELA is English language arts.

Note: The table shows the characteristics for the schools who experienced a principal transition during two time periods (2015/16 to 2018/19 and 2017/18 to 2018/19) relative to all schools in the district. Schools are only counted as experiencing a transition if they are present in the data in both years, the principal changes, and the school has Insight survey data in both years.

Source: Authors' analyses based on administrative data and survey data provided by the District of Columbia Public Schools, 2015/16 to 2018/19.

To estimate the principal value-added model, we regress each relevant school-level survey outcome on a vector of school indicators (fixed effects), a vector of principal indicators, and a vector of time-varying school demographic characteristics—including total enrollment, the percent of students directly certifying for free and reduced-price lunch, and the percent of students by race/ethnicity—following an approach used in the principal value-added literature (Chiang et al., 2016; Grissom et al., 2015). To estimate the models, we use a sample of all schools that experienced a leadership transition from 2015/16 to 2018/19. We standardize each school-level survey outcome relative to the mean across all schools within each year. The school fixed effects adjust for all time-invariant, school-specific factors out of a principal's control that may influence survey outcomes.

The principal fixed effects provide an estimate of each principal's effect on each survey outcome when they move schools. The principal effects are measured in standard deviations of school-level teacher survey measures. The estimates compare principals relative to the set of principals who have served in the same set of schools from 2015/16 to 2018/19. Following Chiang et al. (2016), we define a set as a group of schools in which every school has at least one principal that has served at another school in the group over the analysis period. Because

principals often come into the district and leave without serving at another school, many of the sets used in the analysis only include two principals who served at the same school but did not serve at another school in the district. Following the principal value-added literature, we center all value-added estimates within sets, which implies that the principal value-added estimates are comparisons relative to the average principal in each network. We also adjust the standard errors of the value-added estimates to account for this.

To summarize the principal effects, we present the standard deviation of the principal effects. However, due to measurement error, the variance of the fixed effects is larger than the true variance. As a result, we also calculate and report the true variance of the fixed effects, which is the variance of the fixed effects minus the mean of the squared standard errors of the principal effects (Grissom et al., 2015).

To provide a benchmark of comparison, we also use the same analytical approach to estimate effects on test-based outcomes. To do so, we calculated the percentage of students in each school whose standardized test scores on the PARCC exam were proficient for ELA and for math in each year from 2015/16 to 2018/19; we then averaged these measures to get an average percent proficient for each school in each year. We standardized this measure within each year. Using this standardized measure as the outcome of interest, we calculated principal effects on percent proficient, and we reported the same standard deviations that we used for the survey measure outcomes to provide a comparison.

Research question 4. What is the relationship between School Leader IMPACT components and relevant staff survey measures of the school environment?

To understand the relationship between the existing School Leader IMPACT evaluation system and the school-level averages of the teacher survey measures for each year, the study team calculated pairwise correlations. When interpreting the pairwise correlations, we define small correlations as 0.1 to less than 0.3, moderate as 0.3 to less than 0.5, and large as 0.5 or higher. The School Leader IMPACT evaluation system consists of a final evaluation score, which is an equal average of a Student Outcome Goals score and a Leadership Framework score. The Leadership Framework score consists of a midyear score and an end-of-year score, which are 40 percent and 60 percent of the total Leadership Framework score, respectively. The study team calculated pairwise correlations between each school-level average teacher survey measure and the final score for each year a survey was offered. A handful of principals served in multiple schools in a given year; for these years, we averaged the school survey measures across the two schools in which the principal served. For Panorama, we used data from 2017/18 and 2018/19; for Insight, we used data from 2015/16 to 2018/19. Two Insight domains were only offered in 2017/18 and 2018/19, so these correlations are also limited to these years. We also calculated pairwise correlations between each school-level average teacher survey measure and the Student Outcome Goals score, total Leadership Framework score, and the midyear and end-of-year Leadership Framework scores—pooling estimates across all years.

For the midyear and end-of-year scores, principals are evaluated by their supervisors on six standards. We calculated average scores on the six standards, taking an equal average of the midyear and end-of-year standard scores and calculated pairwise correlations between each school-level average teacher survey measure and six average standard scores. This principal evaluation data containing scores for each standard were only available for 2018/19, so these correlations were limited to one year.

References

- Chiang, H., Lipscomb, S., & Gill, B. (2016). Is school value added indicative of principal quality? *Education Finance and Policy*, 16(3), 283–309.
- Grissom, J., Kalogrides, D., & Loeb, S. (2015). Using student test scores to measure principal performance. *Educational Evaluation and Policy Analysis*, 37(1), 3–28.
- Lipscomb, S., Chiang, H., & Gill, B. (2012). *Value-added estimates for phase 1 of the Pennsylvania teacher and principal evaluation pilot: Full report*. Mathematica Policy Research. <https://eric.ed.gov/?id=ED531795>
- Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, 47(5), 609–612.
- Shiferaw, M. & Gill, B. (2019, December 10). Real-time feedback makes an IMPACT. *RElevant: Viewpoints and Findings from the REL Mid-Atlantic*. <https://ies.ed.gov/ncee/edlabs/regions/midatlantic/app/Blog/Post/26>

Appendix C. Supporting analysis

This appendix presents supporting analyses for each research question.

Research question 1. How do staff survey scale scores vary within and between schools and can they meaningfully differentiate schools from one another?

This section presents results (shown in tables C1 through C3) for research question 1.

Table C1. Intraclass correlations (2017/18, 2018/19)

Survey measure	All years	2017/18	2018/19
Insight			
Academic Opportunity	0.18	0.16	0.21
Diversity, Equity, and Inclusion	0.15	0.14	0.17
Evaluation	0.12	0.13	0.12
Family & Community Engagement	0.18	0.19	0.21
Instructional Planning for Student Growth	0.11	0.12	0.12
Leadership	0.16	0.18	0.18
Learning Environment	0.22	0.23	0.23
Observation & Feedback	0.10	0.10	0.10
Peer Culture	0.14	0.14	0.15
Professional Development	0.11	0.14	0.10
School Operations	0.17	0.17	0.20
Panorama			
Perseverance	0.14	0.13	0.16
Rigorous Expectations	0.04	0.03	0.04
Professional Learning about SEL	0.13	0.14	0.14
Educating All Students	0.03	0.03	0.03
Staff Engagement	0.12	0.12	0.12

SEL is social-emotional learning.

Note: The table shows the proportion of the total variance in each survey scale (Panorama) or domain (Insight) that is attributed to schools, as opposed to teachers.

Source: Authors' analyses based on Insight and Panorama survey data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table C2. School-level rates of favorable responses on the Insight teacher survey (2018/19)

Insight teacher survey	25th percentile	75th percentile	Difference between percentiles
Academic Opportunity	75.0	94.2	19.2
1. Leaders at my school have the necessary content knowledge or content resources to support instruction across subjects.	66.1	88.9	22.8
2. My school implements a rigorous academic curriculum.	81.8	100.0	18.2
3. Students at my school can achieve the academic standards for their grade level.*	70.8	95.7	24.8
4. Students at my school support their answers and explain their thinking.	83.3	100.0	16.7
Diversity, Equity, and Inclusion	67.9	88.1	20.2
5. At my school, my perspective is respected even when it is not aligned with the majority.*	62.5	85.7	23.2
6. My school has a diverse staff.*	72.7	95.5	22.7
7. My school's leaders make promoting diversity and inclusion a high priority.	66.7	90.9	24.2
8. My school's leaders have encouraged me to engage in learning around diversity, equity, and/or inclusion.*	62.5	91.9	29.4
9. People from a diversity of backgrounds and perspectives have fair access to leadership roles at my school.	66.7	88.2	21.6
10. Students at my school value each other's differences (such as gender, race, culture, ability, sexual identity, and learning differences).	61.8	90.9	29.1
11. Teachers and staff at my school discuss how our own identities influence the way we interact with students.*	64.7	88.9	24.2
Evaluation	61.0	81.7	20.8
12. At my school, evaluation ratings are accurate reflections of teacher effectiveness.	47.6	76.2	28.6
13. I agree with the criteria that will be used to evaluate my performance as a teacher.	52.4	76.5	24.1
14. I know the criteria that will be used to evaluate my performance as a teacher.	81.6	95.2	13.7
15. The expectations for effective teaching are clearly defined at my school.	66.7	88.2	21.6
16. The person who evaluates my performance has an accurate perception of my classroom practice.	63.0	85.0	22.0
17. The person who evaluates my performance knows how much progress my students have made this year.	58.8	82.1	23.3
18. The teacher evaluation process helps identify my strengths and weaknesses.	50.0	75.0	25.0
Family & Community Engagement	76.7	95.6	18.9
19. Families at my school regularly receive useful updates about their student's progress.	83.9	100.0	16.1
20. Leaders at my school set clear expectations for family and community engagement.	69.6	92.9	23.3
21. My school encourages and helps families to support student learning at home.	77.3	97.0	19.7
22. My school has established systems that ensure families are well-informed about the school.	75.0	100.0	25.0
23. My school regularly seeks input from students' families.	71.4	94.5	23.1
Instructional Planning for Student Growth	64.6	84.3	19.7
24. An instructional leader at my school regularly reviews student work from my classes.	45.8	72.2	26.4
25. I am satisfied with the support I receive at my school for instructional planning.	60.0	83.3	23.3
26. I collaborate at least weekly with teachers and leaders at my school to improve my instructional plans based on student responses to tasks.	68.9	87.1	18.2
27. I have access to questions, tasks, and assessments that allow me to assess students' understanding of learning goals.	82.1	95.8	13.7
28. I have ready access to content experts with deep subject-area knowledge to support my instructional planning.	61.5	83.3	21.8
29. My school has dedicated time for teachers to analyze student work and/or assessments to plan for future instruction based on student performance.	68.7	88.9	20.2

Table C2. (continued)

Insight teacher survey	25th percentile	75th percentile	Difference between percentiles
Leadership	56.3	80.5	24.3
30. Leaders at my school seek out feedback from teachers.	50.0	79.2	29.2
31. Leaders at my school work hard to retain effective teachers.	50.0	81.3	31.3
32. My school has effective instructional leadership.	58.3	84.2	25.9
33. My school leaders articulate a clear overarching vision that drives priorities, goals, and decisionmaking within the school.	57.7	87.5	29.8
34. My school leaders model the behavior they hope to see across the school community.	57.1	84.6	27.5
35. Teachers understand how our actions contribute to school priorities and goals.	66.7	88.9	22.2
36. When my school leadership commits to a program or priority, they follow through.	50.0	82.8	32.8
Learning Environment	49.3	81.3	32.0
37. Across my school, there are consistent expectations and consequences for student behavior.	31.6	65.0	33.4
38. Interactions between students and adults at my school are respectful.	44.4	85.7	41.3
39. My school is a good place to teach and learn.	66.7	93.3	26.7
40. School leaders promote a safe and productive learning environment in my school.	60.0	88.9	28.9
41. School leaders provide me with the support I need to maintain high standards for student behavior in my classroom.	46.4	77.9	31.5
42. Teachers and leaders at my school immediately address misbehavior in shared school spaces like hallways and the lunch room.	48.4	80.0	31.6
Observation & Feedback	59.8	79.1	19.3
43. Each time I am observed, I get feedback that gives me specific actions to improve my teaching practice.	71.1	87.5	16.4
44. I get enough feedback on my instructional practice.	59.4	83.3	24.0
45. I regularly discuss feedback about my teaching with an instructional leader at my school.	53.3	77.4	24.1
46. My observer consistently follows up to see how I am implementing feedback from our last observation.	45.8	71.0	25.1
47. The feedback I get from being observed helps me improve student outcomes.	68.8	88.5	19.7
48. When I am observed, I get feedback on what I am teaching (the content of the lesson) in addition to how I am teaching.	68.4	84.6	16.2
49. When I get feedback after an observation, I receive support to implement those changes (for example, someone models suggestions for me, or I have time to practice outside of class).	48.1	72.0	23.9
Peer Culture	73.2	90.2	17.0
50. At my school, teachers use a common vocabulary to discuss effective teaching practice.	73.0	93.3	20.4
51. Teachers at my school share a common vision of what effective teaching looks like.	68.3	89.5	21.2
52. The time I spend collaborating with my colleagues is productive.	76.9	92.3	15.4
53. There are many teachers at my school who set an example of what highly effective teaching looks like.	81.8	94.4	12.6
54. There is a low tolerance for ineffective teaching at my school.	61.9	83.3	21.4
Professional Development	59.4	78.6	19.2
55. Coaching and feedback at my school include demonstrations (either live or in video) of what effective teaching of rigorous content looks like.	50.0	73.7	23.7
56. In the past six months, I have practiced teaching techniques with a peer or instructional expert outside my own classroom.	53.1	75.0	21.9
57. In the past six months, someone at my school or district/network has helped me develop new skills or content knowledge that I was able to apply in my own classroom.	66.7	85.7	19.1
58. My school assesses whether the professional development it provides helps teachers improve their instruction.	50.0	76.2	26.2

Table C2. (continued)

	25th percentile	75th percentile	Difference between percentiles
Insight teacher survey			
59. My school is committed to improving my instructional practice.	67.4	88.9	21.5
60. Professional development opportunities at my school are well planned and facilitated.	60.0	84.0	24.0
School Operations	70.6	90.8	20.2
61. Day-to-day operations at my school run smoothly.	60.0	90.9	30.9
62. My school building is clean and well-maintained.	63.6	92.3	28.7
63. My school's systems track student information (such as grading, attendance, and testing) in a way that is useful and up-to-date.	75.0	95.7	20.7
64. Non-academic services for students (such as buses and school meals) are well managed.	84.2	100.0	15.8
65. Procedures at my school maximize the time students spend learning.	59.0	90.5	31.5
66. When I need something at my school, I know who to ask.	78.6	94.4	15.9

Note: For each item on the Insight staff survey, we calculated the percentage of teachers with favorable responses in each school. For each scale, we took the average rate of favorable responses at the school-level across all items in the scale. This table shows the 25th and 75th percentiles of the distribution of the percentage of favorable responses in each school for each item and scale and the difference between the 25th and 75th percentiles.

Source: Authors' analyses based on Insight survey data provided by the District of Columbia Public Schools, 2018/19.

Table C3. School-level rates of favorable responses on the Panorama teacher survey (2018/19)

Panorama teacher survey	25th percentile	75th percentile	Difference between percentiles
Perseverance	32.3	53.2	20.9
1. If your students have a problem while working toward an important goal, how well can they keep working?	25.3	52.6	27.2
2. How often do your students stay focused on the same goal for several months at a time?	22.6	46.9	24.2
3. Some people pursue some of their goals for a long time, and others change their goals frequently. Over the next several years, how likely are your students to continue to pursue one of their current goals?	25.9	47.1	21.1
4. When your students are working on a project that matters a lot to them, how focused can they stay when there are lots of distractions?	35.0	56.5	21.5
5. If your students fail to reach an important goal, how likely are they to try again?	33.3	60.0	26.7
Rigorous Expectations	93.1	98.0	4.8
6. How often do you make your students explain their answers?	87.5	96.8	9.3
7. When students feel like giving up on a difficult task, how likely are you to make them keep trying?	93.0	100.0	7.0
8. How much do you encourage your students to do their best?	96.7	100.0	3.3
9. How often do you take time to make sure your students understand the material?	95.2	100.0	4.8
10. Overall, how high are your expectations of your students?	91.7	100.0	8.3
Professional Learning about SEL	31.2	52.4	21.2
11. In terms of social-emotional learning (SEL) in particular, how supportive has the school been of your growth as a teacher?	35.5	61.7	26.2
12. At your school, how valuable are the social-emotional learning professional development opportunities?	31.7	62.5	30.8
13. When it comes to social-emotional learning, how helpful are your colleagues' ideas for improving your teaching?	39.8	65.0	25.2
14. How often do your social-emotional learning professional development opportunities help you explore new ideas?	21.9	46.9	25.0
15. How relevant have your social-emotional learning professional development opportunities been to the content that you teach?	32.6	54.7	22.1
16. Thinking of social-emotional learning in particular, how much input do you have into individualizing your own professional development opportunities?	22.2	40.0	17.8
17. Overall, how much do you learn about supporting your students' social-emotional learning from the leaders at your school?	26.3	51.5	25.2
Educating All Students	76.9	86.1	9.2
18. How easy do you find interacting with students at your school who are from a different cultural background than your own?	81.8	94.4	12.6
19. How comfortable would you be incorporating new material about people from different backgrounds into your curriculum?	87.0	95.2	8.3
20. How knowledgeable are you regarding where to find resources for working with students who have unique learning needs?	57.1	74.1	16.9
21. If students from different backgrounds struggled to get along in your class, how comfortable would you be intervening?	80.5	92.7	12.2
22. How easy would it be for you to teach a class with groups of students from very different religions from each other?	80.8	89.6	8.8
23. In response to events that might be occurring in the world, how comfortable would you be having conversations about race with your students?	72.3	88.6	16.3
24. How confident are you in ensuring a student who has been socially marginalized feels like a part of your class?	79.7	92.5	12.8

Table C3. (continued)

	25th percentile	75th percentile	Difference between percentiles
Panorama teacher survey			
25. How comfortable would you be working with a student who had difficulty communicating with others because his/her home language is different than the primary language spoken in your classroom and/or school?	72.2	84.8	12.6
26. When a sensitive issue of diversity arises in class, how easily can you think of strategies to address the situation?	66.7	80.0	13.3
Staff Engagement	60.6	74.1	13.5
27. My performance is measured against outcomes and metrics that are clearly explained.	47.5	70.3	22.8
28. I have access to everything that I need in order to perform well at my job.	32.2	60.2	28.0
29. At work, I have the opportunity to put my strengths into practice every day.	63.6	82.5	18.9
30. I regularly receive meaningful recognition for doing my job well.	33.3	55.5	22.2
31. My opinions are taken into account and considered at this job.	40.7	62.5	21.8
32. Working for the purpose or mission of this organization gives me a feeling of accomplishment.	63.3	81.3	18.0
33. My supervisor cares about me as a person.	54.7	79.6	24.9
34. I feel comfortable sharing feedback with my supervisor.	52.6	75.3	22.7
35. My supervisor makes decisions that put students first, even when it is difficult.	48.6	74.5	25.9
36. My colleagues are committed to doing quality work.	74.3	90.2	15.9
37. My colleagues treat me with respect.	83.3	94.4	11.1
38. My team collaborates to reach our goals.	75.6	89.5	13.9
39. On my team, we learn from our failures by trying new things.	64.9	85.1	20.2
40. We regularly take time to celebrate our successes as a team.	46.5	66.0	19.6
41. There is someone at work who encourages my personal progress and development.	70.0	81.4	11.4
42. In my role, there are ongoing opportunities to learn and grow.	56.3	76.3	20.1
43. I enjoy my work at the District of Columbia Public Schools.	66.7	83.1	16.5

Note: For each item on the Panorama teacher survey, we calculated the percentage of teachers with favorable responses in each school. For each scale, we took the average rate of favorable responses at the school-level across all items in the scale. This table shows the 25th and 75th percentiles of the distribution of the percentage of favorable responses in each school for each item and scale and the difference between the 25th and 75th percentiles.

Source: Authors' analyses based on Panorama survey data provided by the District of Columbia Public Schools, 2018/19.

Research question 2. How stable are school-level staff survey measures across years for adjacent years in which the school leader did not change?

This section presents results for research question 2 (table C4). We include the Insight survey data, which are referenced in the main body of the report. However, as noted in appendix B, the analysis for the Panorama survey and two of the Insight domains (Diversity, Equity, and Inclusion; Family & Community Engagement) are only discussed here in the appendix. Due to the small number of schools that experienced a principal transition during the years these survey domains and scales were offered (only 20 schools in 2017/18 and 2018/19), the analyses should be interpreted with caution.

The observed pattern of results for Panorama and the two Insight domains only offered in 2017/18 and 2018/19 is mixed. For Insight, the year-to-year correlations are somewhat weaker for Family & Community Engagement but not for Diversity, Equity, and Inclusion when schools experience a principal transition compared to when they do not. For Panorama, the year-to-year correlations were somewhat weaker when there was a principal transition for Staff Engagement, Educating All Students, and Professional Learning about SEL. Two of these scales are ones that measure perceptions of the school environment (Staff Engagement and Professional Learning about SEL), which principals are more likely to affect. In contrast, the year-to-year correlations are stronger in principal transition years for Rigorous Expectations and Perseverance. This suggests that other school factors that the principals do not affect much were driving the year-to-year correlations for Rigorous Expectations and Perseverance, most likely because these domains reflect district rather than principal priorities. However, because the correlations may be less stable due to small sample size, we are cautious in drawing strong conclusions from these results.

Table C4. Year-to-year correlations for adjacent years in which schools did not experience a principal transition and did experience a principal transition (2015/16 to 2018/19)

Survey measure	Without a principal transition				With a principal transition			
	2015/16 to 2016/17	2016/17 to 2017/18	2017/18 to 2018/19	All years	2015/16 to 2016/17	2016/17 to 2017/18	2017/18 to 2018/19	All years
Insight								
Index	0.719*	0.720*	0.647*	0.687*	0.221	0.430	0.718*	0.454*
Academic Opportunity	0.801*	0.875*	0.760*	0.799*	0.442	0.420	0.857*	0.631*
Evaluation	0.712*	0.795*	0.706*	0.744*	-0.012	0.146	0.616*	0.327*
Instructional Planning for Student Growth	0.733*	0.755*	0.757*	0.725*	0.253	0.382	0.658*	0.472*
Leadership	0.761*	0.772*	0.778*	0.726*	0.147	0.114	0.547*	0.284*
Learning Environment	0.803*	0.787*	0.736*	0.764*	0.487*	0.288	0.754*	0.536*
Observation & Feedback	0.674*	0.715*	0.695*	0.677*	0.200	0.155	0.743*	0.445*
Peer Culture	0.817*	0.777*	0.731*	0.758*	0.411	0.614*	0.703*	0.594*
Professional Development	0.713*	0.714*	0.705*	0.696*	0.183	0.296	0.615*	0.408*
School Operations	0.759*	0.804*	0.698*	0.759*	0.266	0.424	0.660*	0.472*
Observations	83	89	88	260	18	19	20	57
Diversity, Equity, and Inclusion	na	na	0.706*	0.706*	na	na	0.715*	0.715*
Family & Community Engagement	na	na	0.778*	0.778*	na	na	0.666*	0.666*
Observations			88	88			20	20
Panorama								
Perseverance	na	na	0.730*	0.730*	na	na	0.829*	0.829*
Rigorous Expectations	na	na	0.472*	0.472*	na	na	0.842*	0.842*
Professional Learning about SEL	na	na	0.733*	0.733*	na	na	0.651*	0.651*
Educating All Students	na	na	0.652*	0.652*	na	na	0.460*	0.460*
Staff Engagement	na	na	0.729*	0.729*	na	na	0.633*	0.633*
Observations	na	na	91	91	na	na	20	20

na is not available. SEL is social-emotional learning.

* Significant at $p < .05$.

Note: The table shows the year-to-year correlations in school-level average survey measures between adjacent years. The sample used for each correlation only includes schools which did not experience a principal change over the adjacent years. The “all years” column includes adjacent-year pairs for all years that experienced a transition or did not experience a transition. “na” indicates the data are not available for the adjacent years (Panorama was only offered in 2017/18 and 2018/19, and two Insight domains were not offered in 2015/16 or 2016/17). The sample sizes differ for 2018/19 for Panorama and Insight because three schools are missing Insight survey data, but they have data for Panorama.

Source: Authors’ analyses based on Insight and Panorama survey data provided by the District of Columbia Public Schools, 2015/16 to 2018/19.

Research question 3. Which survey scale scores respond more to changes in school leadership?

This section presents results for research question 3 (table C5).

Table C5. Principal effects on teacher survey measures (2015/16 to 2018/19)

Survey measure	Standard deviation of principal effects	True standard deviation	Number of principal effects	Number of principal-year observations
Insight survey				
Index	0.68	0.47	101	195
Academic Opportunity	0.55	0.40	101	195
Evaluation	0.65	0.47	101	195
Instructional Planning for Student Growth	0.67	0.52	101	195
Leadership	0.77	0.61	101	195
Learning Environment	0.61	0.45	101	195
Observation & Feedback	0.66	0.46	101	195
Peer Culture	0.58	0.39	101	195
Professional Development	0.73	0.55	101	195
School Operations	0.59	0.45	101	195
Percent proficient				
Percent proficient (average for ELA and math)	0.14	0.07	101	195

ELA is English language arts.

Note: The first column of this table shows the standard deviation of the principal effect estimates. The second column shows the true standard deviation, which is the variance of the fixed effects minus the mean of the squared standard errors. The estimates for the Insight index and domains are estimated with the set of schools who experienced a principal transition from 2015/16 to 2018/19.

Source: Authors' analyses based on administrative data and survey data provided by the District of Columbia Public Schools, 2015/16 to 2018/19.

Research question 4. What is the relationship between School Leader IMPACT components and relevant staff survey measures of the school environment?

This section presents results for research question 4 (tables C6 through C9).

Table C6. Correlations between staff surveys and principals' total IMPACT score (2015/16 to 2018/19)

Survey measure	School Leader IMPACT scores				
	All years	2015/16	2016/17	2017/18	2018/19
Insight					
Index	0.303*	0.301*	0.412*	0.352*	0.151
Academic Opportunity	0.391*	0.371*	0.517*	0.469*	0.251*
Evaluation	0.245*	0.229*	0.356*	0.314*	0.067
Instructional Planning for Student Growth	0.301*	0.327*	0.411*	0.369*	0.106
Leadership	0.270*	0.213*	0.369*	0.355*	0.147
Learning Environment	0.375*	0.329*	0.493*	0.412*	0.259*
Observation & Feedback	0.199*	0.191	0.337*	0.302*	0.016
Peer Culture	0.258*	0.229*	0.356*	0.316*	0.111
Professional Development	0.322*	0.322*	0.462*	0.323*	0.200*
School Operations	0.369*	0.397*	0.469*	0.370*	0.232*
Observations	418	100	105	107	106
Diversity, Equity, and Inclusion	0.276*	na	na	0.361*	0.198*
Family & Community Engagement	0.307*	na	na	0.358*	0.261*
Observations	213			107	106
Panorama					
Perseverance	0.414*	na	na	0.482*	0.352*
Rigorous Expectations	0.190*	na	na	0.169	0.213*
Professional Learning about SEL	0.282*	na	na	0.291*	0.273*
Educating All Students	0.070	na	na	0.115	0.024
Staff Engagement	0.278*	na	na	0.310*	0.249*
Observations	215	na	na	107	108

Absolute value of the correlation:

0-0.09

0.10–0.29

0.30-0.49

0.50+

na is not available. SEL is social-emotional learning.

*Significant at $p < .05$.

Note: The table shows pairwise correlations between the surveys (Insight survey domains and index scores and the Panorama scale scores) and the District of Columbia Public Schools' total School Leader IMPACT score for the time period noted in the column heading. The color intensity indicates the magnitude of the absolute value of the correlations, using the legend above. Nine Insight domains and the Insight index are available from 2015/16 to 2018/19. Two Insight domains and all Panorama scale scores are only available for 2017/18 and 2018/19.

Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2015/16 to 2018/19.

Table C7. Correlations between staff surveys and principals' IMPACT subscores (2015/16 to 2018/19)

Survey Measure	Weighted Student Outcome Goals score	Weighted Leadership Framework score	Leadership Framework score (cycle 1)	Leadership Framework score (cycle 2)
Insight				
Index	0.173*	0.332*	0.344*	0.392*
Academic Opportunity	0.206*	0.456*	0.502*	0.524*
Evaluation	0.135*	0.277*	0.311*	0.333*
Instructional Planning for Student Growth	0.175*	0.325*	0.329*	0.376*
Leadership	0.144*	0.312*	0.327*	0.356*
Learning Environment	0.214*	0.410*	0.430*	0.469*
Observation & Feedback	0.089	0.258*	0.250*	0.301*
Peer Culture	0.166*	0.383*	0.394*	0.432*
Professional Development	0.160*	0.260*	0.269*	0.319*
School Operations	0.203*	0.417*	0.412*	0.461*
Observations	418	418	415	393
Diversity, Equity, and Inclusion	0.121	0.346*	0.407*	0.398*
Family & Community Engagement	0.165*	0.331*	0.469*	0.462*
Observations	213	213	210	188
Panorama				
Perseverance	0.263*	0.376*	0.485*	0.551*
Rigorous Expectations	0.186*	0.060	0.129	0.195*
Professional Learning about SEL	0.172*	0.270*	0.361*	0.377*
Educating All Students	0.113	-0.057	-0.039	-0.035
Staff Engagement	0.156*	0.291*	0.366*	0.373*
Observations	215	215	212	190

Absolute value of the correlation:

0-0.09

0.10–0.29

0.30-0.49

0.50+

SEL is social-emotional learning.

*Significant at $p < .05$.

Note: The table shows pairwise correlations between the surveys (Insight survey domains and index scores and the Panorama scale scores) and the District of Columbia Public Schools' IMPACT subscores for the time period noted in the column heading. As described in box 1, the District of Columbia Public Schools' School Leader IMPACT final scores are an average of the weighted Student Outcome Goals score and the weighted Leadership Framework score. The weighted Leadership Framework score consists of two Leadership Framework scores from cycle 1 and cycle 2. The color intensity indicates the magnitude of the absolute value of the correlations, using the legend above. Nine Insight domains and the Insight index are available from 2015/16 to 2018/19. Two Insight domains and all Panorama scale scores are only available for 2017/18 and 2018/19.

Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2015/16 to 2018/19.

Table C8. Correlations between staff surveys and principals' Leadership Framework subscores (2018/19)

Survey Measure	Leadership Framework subscore					
	Instruction	Talent	School Culture	Operations	Family & Community	Personal Leadership
Insight						
Index	0.247*	0.176	0.305*	-0.036	0.130	0.160
Academic Opportunity	0.514*	0.396*	0.441*	0.028	0.093	0.126
Diversity, Equity, and Inclusion	0.273*	0.225*	0.349*	0.010	0.182	0.119
Evaluation	0.253*	0.156	0.288* ^b	-0.042	0.043	0.028
Family & Community Engagement	0.359*	0.303*	0.460*	-0.008	0.237*	0.115
Instructional Planning for Student Growth	0.230*	0.218*	0.260*	-0.018	0.073	0.120
Leadership	0.226*	0.154	0.269*	-0.025	0.113	0.126
Learning Environment	0.356*	0.240*	0.365*	0.027	0.170	0.160
Observation & Feedback	0.141	0.124	0.145	0.030	0.024	0.021
Peer Culture	0.323*	0.190	0.323*	0.008	0.105	0.139
Professional Development	0.238*	0.211*	0.259*	-0.021	0.129	0.130
School Operations	0.307*	0.186	0.429*	0.069	0.271*	0.221*
Observations	106	106	106	106	106	106
Panorama						
Perseverance	0.524*	0.357*	0.397*	0.124	0.028	0.083
Rigorous Expectations	0.160	0.042 ^a	0.181	0.111	-0.192*	-0.074
Professional Learning about SEL	0.295*	0.239*	0.264*	0.187	0.119	0.109
Educating All Students	-0.138	-0.104	-0.015 ^a	-0.022	-0.093	-0.021
Staff Engagement	0.250*	0.243*	0.325*	0.057	0.092	0.167
Observations	108	108	108	108	108	108

Absolute value of the correlation:

0-0.09

0.10–0.29

0.30-0.49

0.50+

SEL is social-emotional learning.

*Significant at $p < .05$.

Note: The table shows pairwise correlations between the surveys (Insight survey domains and index scores and the Panorama scale scores) and principals' Leadership Framework subscores—specifically the average of the subscore from cycle 1 and cycle 2. These detailed Leadership Framework component scores were only available for 2018/19. The color intensity indicates the magnitude of the absolute value of the correlations, using the legend above.

Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2018/19.

Table C9. Correlations between Insight and Panorama (2017/18 and 2018/19)

Insight index and domains	Panorama scales				
	Perseverance	Rigorous Expectations	Professional Learning about SEL	Educating All Students	Staff Engagement
Index	0.755	0.450	0.715	0.184	0.765
Academic Opportunity	0.574	0.261	0.670	0.142	0.762
Diversity, Equity, and Inclusion	0.519	0.378	0.676	0.286	0.767
Evaluation	0.627	0.309	0.693	0.105	0.763
Family & Community Engagement	0.521	0.384	0.678	0.300	0.758
Instructional Planning for Student Growth	0.495	0.310	0.714	0.204	0.808
Leadership	0.666	0.356	0.682	0.201	0.754
Learning Environment	0.422	0.398	0.624	0.369	0.711
Observation & Feedback	0.480	0.348	0.696	0.287	0.759
Peer Culture	0.599	0.392	0.701	0.261	0.775
Professional Development	0.546	0.260	0.666	0.145	0.732
School Operations	0.540	0.368	0.680	0.262	0.766
Observations	224	224	224	224	224

Absolute value of the correlation:

0-0.09

0.10–0.29

0.30-0.49

0.50+

SEL is social-emotional learning.

*Significant at $p < .05$.

Note: The table shows pairwise correlations between the Insight survey domains and index scores and the Panorama scale scores. The color intensity indicates the magnitude of the absolute value of the correlations, using the legend above.

Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2017/17 to 2018/19.

Appendix D. Nonresponse bias analysis

This appendix provides a detailed description of response rates and the nonresponse bias analysis. It also includes a sensitivity check (for the main results) in which we apply survey nonresponse weights.

Analyzing response rates

For the Insight and Panorama teacher surveys, not all teachers eligible to take the surveys did so, and some who completed the survey may not have responded to every item. If the teachers who responded to the survey had responses that were not representative of all teachers in their school, the school-level averages of the survey measures could be misleading. For example, if teachers who have more rigorous expectations for students were more likely to respond to the survey, the school-level score for Panorama's Rigorous Expectation scale may appear higher than it would be if all teachers had responded. To address this, the study examined response rates, checked for evidence of nonresponse bias, and applied weights to adjust for nonresponse, as needed.

Response rates. The following are response rates for each survey measure and sample used in the study for the Panorama (table D1) and Insight (table D2) teacher surveys:

- Unit response rate, which is the number of respondents who took the survey out of those who were eligible to take the survey.
- Scale response rate, which is the number of respondents who had all non-missing values for the items on a given scale out of those who took the survey.
- Overall response rate, which is the number of respondents who had all non-missing values for the items on a given scale out of those who were eligible to take the survey.

For the district as a whole, we report the unit response rate. For each scale, we report the scale response rate and overall response rate. For each type of response rate (unit, scale, and overall), we show the average for the district and we calculate the response rate for each school and report the 25th and 75th percentiles of the school distributions to describe the variation in response rates across schools.

For both Insight and Panorama, we calculate response rates for 2017/18 and 2018/19. Although we use Insight data for 2015/16 and 2016/17 in the analyses in this study, we do not have access to teacher-level responses for Insight in these two years and cannot do a nonresponse bias analysis. However, we do have school-level unit response rates that were provided by the district for 2015/16 and 2016/17. For these, we calculate the 25th and 75th percentiles of the distribution of school unit responses in each year (table D3).

To identify the set of staff eligible to take the survey, the District of Columbia Public Schools (DCPS) provided us with a dataset of the instructional staff eligible to take Panorama in 2017/18 and 2018/19 and teachers eligible to take Insight in 2017/18 and 2018/19. (The instructional staff who take the Panorama survey include teachers and other instructional staff—such as coaches, aides, and teacher residency fellows—while Insight is only offered to teachers. To simplify, we use the term teacher to denote the eligible population for either survey.) Because the Insight roster file of eligible teachers does not include race/ethnicity, gender, and years of experience, we use the teacher characteristics listed in the Panorama roster and limit the Insight roster used in the nonresponse analysis to those teachers that were in the Panorama roster. Less than 0.06 percent of teachers in the Insight roster were not found in the Panorama roster (32 out of 3,836 in 2017/18 and 15 of 4,170 in 2018/19).

The unit response rate for Panorama was 60 percent in 2017/18 and 77 percent in 2018/19 (see table D1). The 25th percentile and 75th percentile of the school-level unit response rates were 45 percent and 76 percent, respectively, in 2017/18 and 68 percent and 88 percent, respectively, in 2018/19. The scale response rates were 93 percent or higher, with most of the teachers who responded to at least one item on the survey responding to

all the items in each scale. The overall response rates for each Panorama scale ranged from 56 to 58 percent in 2017/18 and 72 to 75 percent in 2018/19, which were similar to the unit response rates in each year.

The unit response rate for Insight was 74 percent in 2017/18 and 68 percent in 2018/19 (see table D2). The interquartile range of the school-level response rates was 26 percentage points in 2017/18 and 28 percentage points in 2018/19. Like the Panorama survey, the scale response rates for Insight were high (92 percent or higher), with most of the teachers who responded to at least one item on the survey responding to all the items in each scale. The overall response rates for each Insight domain ranged from 69 to 72 percent in 2017/18 and 63 to 66 percent in 2018/19, which were similar to the unit response rates in each year. We could not calculate teacher-level response rates for Insight in 2015/16 and 2016/17 because we did not have teacher-level survey response data or the roster of eligible teachers. However, we do have the school-level unit response rates. The interquartile range was 22 percentage points in 2015/16 and 26 percentage points in 2016/17 (see table D3), which is similar to the interquartile ranges of the unit response rates in 2017/18 and 2018/19.

Table D1. Panorama teacher survey response rates (2017/18, 2018/19)

Survey measure	2017/18			2018/19		
	All teachers	School-level rates		All teachers	School-level rates	
		25th percentile	75th percentile		25th percentile	75th percentile
Unit response rate	60.0	45.2	75.8	76.9	68.2	88.0
Perseverance						
Scale response rate	96.3	93.8	100.0	96.6	94.0	100.0
Overall response rate	57.8	42.3	73.4	74.4	65.5	84.0
Rigorous Expectations						
Scale response rate	97.0	94.4	100.0	97.9	96.2	100.0
Overall response rate	58.3	44.7	73.9	75.4	66.7	85.7
Professional Learning about SEL						
Scale response rate	95.2	92.0	100.0	96.3	93.5	100.0
Overall response rate	57.2	42.4	71.4	74.1	65.1	84.4
Educating All Students						
Scale response rate	95.3	92.3	100.0	96.1	93.8	100.0
Overall response rate	57.2	42.0	71.4	73.9	64.9	84.4
Staff Engagement						
Scale response rate	92.8	89.5	96.6	93.9	91.7	97.2
Overall response rate	55.7	41.8	71.4	72.2	63.0	83.1
Observations	5,171	118	118	4,051	117	117

SEL is social-emotional learning.

Note: The table shows the unit, scale, and overall response rates to the teacher surveys for the Panorama teacher survey for each survey scale used in the study (described in appendix B).

Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table D2. Insight teacher survey response rates (2017/18, 2018/19)

Survey measure	2017/18			2018/19		
	All teachers	School-level rates		All teachers	School-level rates	
		25th percentile	75th percentile		25th percentile	75th percentile
Unit response rate	73.7	64.3	89.7	68.0	55.0	83.3
Academic Opportunity						
Scale response rate	94.4	90.8	100.0	96.1	93.8	100.0
Overall response rate	69.6	57.7	86.2	65.4	52.0	80.0
Diversity, Equity, and Inclusion						
Scale response rate	93.1	88.2	97.6	92.4	88.9	97.1
Overall response rate	68.6	55.2	83.3	62.8	46.9	77.8
Evaluation						
Scale response rate	95.1	91.7	100.0	95.4	92.9	100.0
Overall response rate	70.1	58.8	86.4	64.9	50.0	80.0
Family & Community Engagement						
Scale response rate	93.2	88.2	97.4	93.1	89.7	100.0
Overall response rate	68.6	56.0	84.0	63.3	46.7	77.8
Instructional Planning for Student Growth						
Scale response rate	97.7	96.2	100.0	95.9	92.9	100.0
Overall response rate	72.0	62.1	87.1	65.2	51.6	80.0
Leadership						
Scale response rate	92.7	88.9	97.6	95.4	91.7	100.0
Overall response rate	68.3	55.6	85.0	64.9	51.4	79.5
Learning Environment						
Scale response rate	95.3	92.5	100.0	96.0	92.6	100.0
Overall response rate	70.2	60.0	86.4	65.3	51.6	79.5
Observation & Feedback						
Scale response rate	94.0	91.3	100.0	95.7	93.8	100.0
Overall response rate	69.3	58.8	84.5	65.1	52.0	80.0
Peer Culture						
Scale response rate	96.6	93.5	100.0	96.3	94.7	100.0
Overall response rate	71.2	60.8	86.8	65.5	50.0	80.0
Professional Development						
Scale response rate	97.8	95.4	100.0	96.4	93.8	100.0
Overall response rate	72.0	60.8	87.1	65.6	52.0	80.6
School Operations						
Scale response rate	92.7	87.5	97.4	92.7	89.5	97.1
Overall response rate	68.3	56.9	84.0	63.0	48.0	78.3
Observations	3,869	115	115	3,989	115	115

Note: The table shows the unit, scale, and overall response rates to the Insight teacher survey for each domain scale used in the study (described in appendix B).

Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table D3. Insight teacher survey response rates (2015/16, 2016/17)

Measure	2015/16			2016/17		
	All teachers	School-level rates		All teachers	School-level rates	
		25th percentile	75th percentile		25th percentile	75th percentile
Unit response rate	na	59.3	80.8	na	64.3	90.0
Observations	na	105	105	na	113	113

na is not available.

Note: The table shows the 25th and 75th percentiles of the school-level unit response rates to the Panorama teacher survey in 2015/16 and 2016/17.

Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2015/16, 2016/17.

Nonresponse bias analysis

Because overall response rates were less than 85 percent (the standard set in the National Center on Education and the Economy [NCEE] guidance), the study team conducted a nonresponse bias analysis to examine the potential for nonresponse bias to affect the results. The analysis is composed of two parts. First, we examine how the characteristics of teachers who responded to the survey differ from the characteristics of all eligible teachers in each year. The teacher characteristics available for these analyses include gender, level of experience, and race/ethnicity. For both the Panorama Insight surveys, less than 8 percent of teachers in each year are missing any of these variables.

The study team calculated the differences in teacher characteristics in standard deviation units between teachers who responded to the survey and teachers in the total eligible sample (table D4 and table D5, respectively). For Insight, the absolute values of the differences were not greater than 0.05 in 2017/18 but were greater for the share of teachers who are Black (-0.08) and the share of teachers who are White (0.07) in 2018/19 (see table D4). For Panorama, the absolute value of the difference between the teachers who responded to the survey and all eligible teachers was larger than 0.05 standard deviations for the share of teachers who are Black (-0.08 for 2017/18) and the share who are White (0.10 in 2017/18 and 0.06 in 2018/19) (see table D5).

Second, we examine the relationship between each survey measure and teacher characteristics to understand if teachers with different characteristics respond differently (tables D6 and D7). For Insight, the absolute value of the correlations between the Insight domains and the following teacher characteristics are greater than 0.1 for at least one domain: 0–1 years of experience, greater than 10 years of experience, and White. However, none of the correlations are larger than 0.15. For Panorama, the absolute value of the correlations between the Panorama scales and the following teacher characteristics were greater than 0.1 for at least one scale: 0–1 years of experience, greater than 10 years of experience, Black, and White. The correlations were never greater than 0.2.

Given that the differences between the teacher characteristics of the sample of eligible teachers and teachers who responded to the Insight survey were never larger than 0.05 standard deviations in 2017/18, it is not necessary to adjust for nonresponse on the Insight survey for that year, according to NCEE guidance. For the Panorama survey and the Insight survey in 2018/19, there were differences in the share of teachers who were White and the share who were Black that were larger than 0.05 standard deviations. Although the correlations between White or Black and the Panorama scales were never greater than 0.2 and the correlations between White or Black and the Insight domains were never greater than 0.14, it is possible that these differences could induce bias.

Table D4. Differences between Insight respondents and the eligible sample of teachers (2017/18, 2018/19)

Scale	Mean for survey sample		Mean (and standard deviation) for original study sample		Difference in means in standard deviation units	
	2017/18	2018/19	2017/18	2018/19	2017/18	2018/19
Female	0.76	0.76	0.75 (0.43)	0.75 (0.43)	0.03	0.03
Teacher's years of experience						
0–1 years	0.24	0.20	0.23 (0.42)	0.21 (0.41)	0.00	-0.03
2–3 years	0.23	0.22	0.23 (0.42)	0.21 (0.41)	0.01	0.02
4–5 years	0.12	0.16	0.12 (0.33)	0.15 (0.36)	0.00	0.01
6–10 years	0.17	0.18	0.17 (0.38)	0.18 (0.38)	0.01	-0.01
>10 years	0.24	0.25	0.24 (0.43)	0.25 (0.43)	-0.02	0.00
Teacher's race/ethnicity						
Black	0.48	0.47	0.50 (0.50)	0.51 (0.50)	-0.05	-0.08
Hispanic	0.07	0.07	0.07 (0.25)	0.07 (0.26)	0.01	0.00
White	0.34	0.35	0.32 (0.47)	0.31 (0.46)	0.05	0.07
Other	0.04	0.04	0.04 (0.20)	0.04 (0.19)	0.01	0.01
Not reported	0.06	0.07	0.07 (0.25)	0.07 (0.25)	-0.01	0.01
Observations	2,848– 2,851	2,710– 2,713	3,864–3,869	3,985–3,989		

Note: The table shows the mean teacher characteristics for teachers who responded to at least one item on the Insight survey in each year and those that were eligible to respond to the survey. The difference columns show the difference between the two means in each year in standard deviation units. The number of observations varies slightly within columns due to some teachers missing data, so ranges are reported for each column.

Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table D5. Differences between Panorama respondents and the eligible sample of teachers (2017/18, 2018/19)

Scale	Mean for survey sample		Mean (and standard deviation) for original study sample		Difference in means in standard deviation units	
	2017/18	2018/19	2017/18	2018/19	2017/18	2018/19
Female	0.79	0.75	0.77 (0.42)	0.77 (0.43)	0.05	0.05
Teacher's years of experience						
0–1 years	0.24	0.21	0.24 (0.43)	0.20 (0.41)	-0.01	-0.02
2–3 years	0.23	0.20	0.22 (0.42)	0.20 (0.40)	0.01	0.00
4–5 years	0.11	0.15	0.12 (0.32)	0.16 (0.36)	-0.01	0.01
6–10 years	0.17	0.18	0.17 (0.38)	0.18 (0.38)	0.01	0.01
>10 years	0.24	0.25	0.24 (0.43)	0.25 (0.43)	0.00	0.00
Teacher's race/ethnicity						
Black	0.50	0.50	0.54 (0.50)	0.48 (0.50)	-0.08	-0.05
Hispanic	0.07	0.07	0.07 (0.26)	0.08 (0.26)	0.01	0.01
White	0.32	0.31	0.27 (0.45)	0.34 (0.46)	0.10	0.06
Other	0.04	0.04	0.04 (0.19)	0.04 (0.19)	0.01	0.00
Not reported	0.07	0.07	0.08 (0.26)	0.07 (0.25)	-0.03	0.00
Observations	3,102– 3,105	3,117	5,166–5,171	4,051		

Note: The table shows the mean teacher characteristics for teachers who responded to at least one item on the Panorama survey in each year and those that were eligible to respond to the survey. The difference columns show the difference between the two means in each year in standard deviation units. The number of observations varies slightly within columns in 2017/18 due to some teachers missing data, so ranges are reported in the 2017/18 columns.

Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table D6. Correlations between teacher characteristics and Insight domains (2017/18, 2018/19)

Scale	Female	Years of experience					Race/ethnicity				
		0–1 years	2–3 years	4–5 years	6–10 years	>10 years	Black	Hispanic	White	Other	Not reported
Academic Opportunity											
2017/18	0.06*	-0.07*	-0.02	-0.02	-0.01	0.12*	0.00	0.02	-0.02	0.01	-0.00
2018/19	0.04	-0.12*	-0.04*	0.01	0.01	0.13*	-0.03	0.04	0.02	-0.01	-0.00
DEI											
2017/18	-0.02	-0.05*	-0.02	-0.02	-0.03	0.11*	-0.02	0.02	0.00	0.03	-0.01
2018/19	-0.06*	-0.06*	-0.06*	0.00	-0.00	0.11*	-0.05*	0.07*	-0.00	0.02	0.01
Evaluation											
2017/18	-0.00	-0.05*	-0.04*	0.01	0.01	0.08*	0.04	-0.01	-0.05*	0.04*	-0.01
2018/19	-0.02	-0.12*	-0.05*	0.02	0.04*	0.11*	0.01	0.04	-0.07*	0.05*	0.03
Family & Community Engagement											
2017/18	0.03	-0.05*	-0.01	-0.02	-0.03	0.10*	0.03	0.02	-0.05*	0.04	-0.01
2018/19	0.02	-0.07*	-0.07*	0.00	-0.00	0.13*	0.03	0.05*	-0.06*	0.04	-0.02
Instructional Planning											
2017/18	-0.02	-0.05*	-0.05*	-0.04*	-0.01	0.14*	0.08*	0.02	-0.11*	0.02	0.02
2018/19	-0.02	-0.08*	-0.05*	-0.02	-0.01	0.15*	0.09*	0.04	-0.12*	0.02	0.00
Leadership											
2017/18	-0.01	-0.03	-0.06*	-0.03	-0.01	0.12*	0.07*	0.02	-0.09*	0.02	-0.00
2018/19	-0.03	-0.07*	-0.06*	-0.02	-0.01	0.15*	0.07*	0.05*	-0.10*	0.02	0.00
Learning Environment											
2017/18	0.00	-0.08*	-0.05*	-0.03	0.01	0.14*	0.02	0.06*	-0.07*	0.05*	0.00
2018/19	-0.02	-0.09*	-0.07*	-0.01	0.01	0.15*	0.01	0.06*	-0.07*	0.05*	0.01
Observation & Feedback											
2017/18	-0.05*	-0.00	-0.06*	-0.04*	-0.03	0.12*	0.10*	0.01	-0.14*	0.04*	0.02
2018/19	-0.06*	-0.04*	-0.06*	-0.01	-0.03	0.13*	0.10*	0.05*	-0.14*	0.02	0.01
Peer Culture											
2017/18	0.00	-0.04*	-0.04*	-0.04*	-0.02	0.13*	0.05*	-0.02	-0.06*	0.03	-0.01
2018/19	-0.00	-0.06*	-0.07*	-0.03	-0.00	0.15*	0.04	0.03	-0.06*	0.03	-0.01
Professional Development											
2017/18	-0.03	-0.03	-0.05*	-0.05*	-0.02	0.14*	0.08*	0.02	-0.12*	0.03	0.00
2018/19	-0.03	-0.05*	-0.06*	-0.03	-0.02	0.14*	0.08*	0.03	-0.12*	0.02	0.01
School Operations											
2017/18	-0.01	-0.07*	-0.03	-0.01	-0.01	0.13*	0.07*	0.01	-0.10*	0.06*	-0.01
2018/19	-0.02	-0.09*	-0.04*	-0.01	0.01	0.13*	0.06*	0.03	-0.10*	0.04	0.01

DEI is Diversity, Equity, and Inclusion.

*Significant at $p < .05$.

Note: The table shows the correlations between teacher characteristics and each of the Insight domains for 2017/18 and 2018/19.

Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table D7. Correlations between teacher characteristics and Panorama survey scales (2017/18, 2018/19)

Scale	Female	Years of experience					Race/ethnicity				
		0–1 years	2–3 years	4–5 years	6–10 years	>10 years	Black	Hispanic	White	Other	Not reported
Perseverance											
2017/18	0.05*	-0.07*	-0.03	-0.02	0.02	0.10*	0.04*	0.06*	-0.09*	0.03	0.02
2018/19	0.06*	-0.08*	-0.02	-0.01	0.01	0.09*	0.00	0.06*	-0.05*	0.02	0.01
Rigorous Expectations											
2017/18	0.10*	-0.11*	-0.01	0.02	0.02	0.08*	0.11*	-0.01	-0.11*	-0.04*	0.03
2018/19	0.10*	-0.09*	-0.02	0.01	0.01	0.08*	0.10*	0.03	-0.13*	-0.02	0.01
Professional Learning about SEL											
2017/18	0.01	-0.03	-0.04*	-0.05*	-0.03	0.13*	0.08*	0.07*	-0.16*	0.05*	0.03
2018/19	-0.03	-0.04*	-0.05*	-0.01	-0.03	0.12*	0.02	0.05*	-0.08*	0.04*	0.02
Educating All Students											
2017/18	-0.02	-0.06*	-0.03	0.00	0.00	0.08*	0.19*	-0.01	-0.20*	-0.03	0.01
2018/19	-0.01	-0.04*	-0.04*	0.00	0.01	0.07*	0.20*	-0.01	-0.20*	-0.05*	0.03
Staff Engagement											
2017/18	0.01	-0.06*	-0.05*	-0.01	-0.01	0.13*	0.04*	0.04*	-0.08*	0.05*	-0.00
2018/19	-0.02	-0.06*	-0.07*	-0.01	-0.01	0.14*	0.02	0.01	-0.06*	0.04*	0.03

SEL is social-emotional learning.

*Significant at $p < .05$.

Note: The table shows the correlations between teacher characteristics and each of the Panorama survey scales.

Source: Authors' analyses based on survey and administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Nonresponse weights

Given the possibility of nonresponse bias on the Panorama survey, the study team created nonresponse weights for teachers for the Panorama survey. Nonresponse weights help to make the survey responses more representative of the eligible population of teachers by giving more weight to responding teachers who are part of a group that is less likely to have responded to the survey. Because we use school-level averages of the survey measures in this report, we construct the nonresponse weights within schools using a weighting class approach described below:

- 1) The teacher characteristics that are imbalanced between the group of eligible teachers and the group of responding teachers are the race/ethnicity variables, specifically Black and White. To address this, we create three groups within each school: Black teachers, White teachers, and all other teachers.
- 2) For each group in each school, we calculate nonresponse weights equal to the number of group members in the eligible population of teachers in the school over the number who responded to the survey, which we define as answering at least one item on the survey.

We then re-create the major tables from appendix C, applying nonresponse weights when constructing school-level averages of the Panorama scales and Insight domains. Below is a summary of the findings by research question:

Research question 1: As shown below, for both the Insight (table D8) and Panorama (table D9) surveys, there is little difference between the weighted and unweighted results. The main conclusion holds with the weights or

without them: there are meaningful differences (greater than 15 percentage points) across schools in the share of teachers who have favorable responses to all of the Insight domains and two of the Panorama scales (Perseverance and Professional Learning about SEL). One Panorama scale (Staff Engagement) has moderate differences (between 10 and 15 percentage points) between schools in the 25th and 75th percentiles of the distribution and two (Educating All Students and Rigorous Expectations) had modest differences (less than 10 percentage points).

Table D8. School-level rates of favorable responses on the Insight teacher survey with nonresponse weights applied (2018/19)

Survey measure	Weighted			Unweighted		
	25th percentile	75th percentile	Difference between percentiles	25th percentile	75th percentile	Difference between percentiles
Academic Opportunity	75.7	93.1	17.4	75.0	94.2	19.2
Diversity, Equity, and Inclusion	66.1	87.9	21.8	67.9	88.1	20.2
Evaluation	61.2	82.3	21.2	61.0	81.7	20.8
Family & Community Engagement	76.3	94.8	18.5	76.7	95.6	18.9
Instructional Planning for Student Growth	64.2	83.8	19.7	64.6	84.3	19.7
Leadership	59.2	82.4	23.2	56.3	80.5	24.3
Learning Environment	51.5	80.9	29.4	49.3	81.3	32.0
Observation & Feedback	60.5	78.7	18.2	59.8	79.1	19.3
Peer Culture	72.4	90.9	18.5	73.2	90.2	17.0
Professional Development	60.9	80.4	19.6	59.4	78.6	19.2
School Operations	72.5	90.2	17.7	70.6	90.8	20.2

Note: For each item on the Insight teacher survey, we calculated the percentage of teachers with favorable responses in each school. For each scale, we took the average rate of favorable responses at the school-level across all items in the scale. This table shows the 25th and 75th percentiles of the distribution of the percentage of favorable responses in each school for each item and scale. The first three columns have nonresponse weights applied when constructing school-level averages, while the last three do not.

Source: Authors' analyses based on Insight survey data provided by the District of Columbia Public Schools, 2018/19.

Table D9. School-level rates of favorable responses on the Panorama teacher survey with nonresponse weights applied (2018/19)

Survey measure	Weighted			Unweighted		
	25th percentile	75th percentile	Difference between percentiles	25th percentile	75th percentile	Difference between percentiles
Perseverance	31.4	51.8	20.4	32.3	53.2	20.9
Rigorous Expectations	92.5	97.6	5.1	93.1	98.0	4.8
Professional Learning about SEL	31.0	51.8	20.8	31.2	52.4	21.2
Educating All Students	76.5	85.5	9.0	76.9	86.1	9.2
Staff Engagement	60.1	73.1	13.0	60.6	74.1	13.5

SEL is social-emotional learning.

Note: For each item on the Panorama teacher survey, we calculated the percentage of teachers with favorable responses in each school. For each scale, we took the average rate of favorable responses at the school-level across all items in the scale. This table shows the 25th and 75th percentiles of the distribution of the percentage of favorable responses in each school for each item and scale. The first three columns have nonresponse weights applied when constructing school-level averages, while the last three do not.

Source: Authors' analyses based on Panorama survey data provided by the District of Columbia Public Schools, 2018/19.

Research question 2: For the Insight (table D10) and Panorama (table D11) surveys, the year-to-year correlations using unweighted data or when applying nonresponse weights are typically similar. The average of the absolute value of the difference between the weighted and unweighted correlations in tables D10 and D11 is 0.07.

Because the study team only has teacher-level data for 2017/18 and 2018/19, we cannot construct nonresponse weights for 2015/16 or 2016/17. This limits this robustness analysis to only one year of principal transitions: 2017/18 and 2018/19. For this set of adjacent years, the year-to-year correlations did not show the same pattern in the unweighted data that the other years did. Specifically, the year-to-year correlations did not consistently decline in years when a principal transitioned compared to years in which they did not transition. These findings were replicated in the weighted data.

Table D10. Year-to-year correlations for Insight for adjacent years in which schools did not experience a principal transition and did experience a principal transition with nonresponse weights applied (2017/18 to 2018/19)

Survey measure	Without a principal transition	With a principal transition
	2017/18 to 2018/19	2017/18 to 2018/19
Weighted		
Academic Opportunity	0.662*	0.879*
Evaluation	0.705*	0.610*
Instructional Planning for Student Growth	0.611*	0.691*
Leadership	0.607*	0.544*
Learning Environment	0.673*	0.767*
Observation & Feedback	0.645*	0.683*
Peer Culture	0.598*	0.738*
Professional Development	0.559*	0.609*
School Operations	0.631*	0.699*
Diversity, Equity, and Inclusion	0.657*	0.879*
Family & Community Engagement	0.666*	0.610*
Unweighted		
Academic Opportunity	0.647*	0.857*
Evaluation	0.760*	0.616*
Instructional Planning for Student Growth	0.706*	0.658*
Leadership	0.757*	0.547*
Learning Environment	0.778*	0.754*
Observation & Feedback	0.736*	0.743*
Peer Culture	0.695*	0.703*
Professional Development	0.731*	0.615*
School Operations	0.698*	0.660*
Diversity, Equity, and Inclusion	0.706*	0.715*
Family & Community Engagement	0.778*	0.666*

*Significant at $p < .05$.

Note: The table shows the year-to-year correlations in school-level average survey measures between adjacent years in which a principal transition occurred and in which it did not. The top panel shows the results when the school-level average survey measures are constructed using nonresponse weights, while the bottom panel shows the unweighted results. The sample used for each correlation only includes schools which did not experience a principal change over the adjacent years.

Source: Authors' analyses based on survey data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table D11. Year-to-year correlations for Panorama for adjacent years in which schools did not experience a principal transition and did experience a principal transition with nonresponse weights applied (2015/16 to 2018/19)

Survey measure	Without a principal transition	With a principal transition
	2017/18 to 2018/19	2017/18 to 2018/19
Weighted		
Perseverance	0.715*	0.778*
Rigorous Expectations	0.255*	0.685*
Professional Learning about SEL	0.706*	0.632*
Educating All Students	0.411*	0.386
Staff Engagement	0.665*	0.628*
Unweighted		
Perseverance	0.730*	0.829*
Rigorous Expectations	0.472*	0.842*
Professional Learning about SEL	0.733*	0.651*
Educating All Students	0.652*	0.460*
Staff Engagement	0.729*	0.633*

SEL is social-emotional learning.

*Significant at $p < .05$.

Note: The table shows the year-to-year correlations in school-level average survey measures between adjacent years in which a principal transition occurred and in which it did not. The top panel shows the results when the school-level average survey measures are constructed using nonresponse weights, while the bottom panel shows the unweighted results. The sample used for each correlation only includes schools that did not experience a principal change over the adjacent years.

Source: Authors' analyses based on survey data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Research question 3: We estimate the principal effects model with unweighted data for just the two years of data for which we can construct nonresponse weights (2017/18 and 2018/19) (table D12). Using all four years of data, the unweighted true standard deviations of the principal effects range from 0.39 to 0.61. With just the two years of data, the unweighted true standard deviations range from 0.40 to 0.68, which is very similar. The weighted results using just the two years of data have a slightly wider range (0.28 to 0.79), but the main conclusions still hold. The true standard deviations of the principal effects on the Insight survey measures are still much larger than the effects on percent proficient in math and English language arts (0.07). The domain with the largest standard deviation (Leadership) is also the same using weighted or unweighted data.

Table D12. Principal effects on teacher survey measures (2017/18 to 2018/19)

Survey measure	Weighted		Unweighted	
	Standard deviation of principal effects	True standard deviation	Standard deviation of principal effects	True standard deviation
Insight survey				
Academic Opportunity	0.38	0.30	0.44	0.44
Evaluation	0.77	0.76	0.58	0.55
Instructional Planning for Student Growth	0.56	0.32	0.56	0.50
Leadership	0.81	0.79	0.68	0.67
Learning Environment	0.41	0.28	0.41	0.23
Observation & Feedback	0.50	0.46	0.40	0.40
Peer Culture	0.55	0.55	0.61	0.61
Professional Development	0.76	0.57	0.60	0.56
School Operations	0.47	0.47	0.58	0.54

Note: The first two columns use nonresponse weights to calculate school-level averages for each survey measure, while the last two columns are unweighted. Within the weighted and unweighted section, the first column shows the standard deviation of the principal effect estimates. The second column shows the true standard deviation, which is the variance of the fixed effects minus the mean of the squared standard errors. The estimates for the Insight domains are estimated with the set of schools that experienced a principal transition from 2017/18 to 2018/19, because weights cannot be estimated for 2015/16 and 2016/17. The Insight index is missing because weights cannot be estimated for the index.

Source: Authors' analyses based on administrative data and survey data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Research question 4: The main findings are consistent when using weighted or unweighted data. All of the Insight domains and four Panorama scales are significantly correlated with the overall principal IMPACT score when using or not using nonresponse weights (tables D13 and D14). All of the Insight domains and the three Panorama scales that are most strongly correlated with the final School IMPACT score are more correlated with the Leadership Framework scores than the Student Outcome Goals scores when using or not using nonresponse weights (tables D15 and D16). The Insight domains and the three Panorama scales that are significantly correlated with the Leadership Framework scores are most strongly correlated with the following standards within the Leadership Framework—Instruction, Talent, and School Culture—when using or not using nonresponse weights (tables D17 and D18).

Table D13. Correlations between Insight domains and principals' total IMPACT score with nonresponse weights applied (2017/18 to 2018/19)

Survey measure	Weighted both years	Weighted 2017/18	Weighted 2018/19	Unweighted both years	Unweighted 2017/18	Unweighted 2018/19
Academic Opportunity	0.377*	0.489*	0.258*	0.391*	0.469*	0.251*
Evaluation	0.202*	0.344*	0.048	0.245*	0.314*	0.067
Instructional Planning for Student Growth	0.249*	0.387*	0.102	0.301*	0.369*	0.106
Leadership	0.265*	0.379*	0.145	0.270*	0.355*	0.147
Learning Environment	0.357*	0.429*	0.280*	0.375*	0.412*	0.259*
Observation & Feedback	0.182*	0.324*	0.034	0.199*	0.302*	0.016
Peer Culture	0.283*	0.347*	0.221*	0.258*	0.316*	0.111
Professional Development	0.229*	0.340*	0.108	0.322*	0.323*	0.200*
Diversity, Equity, and Inclusion	0.280*	0.387*	0.173	0.276*	0.361*	0.198*
Family & Community Engagement	0.320*	0.408*	0.247*	0.307*	0.358*	0.261*

Absolute value of the correlation:

0-0.09

0.10–0.29

0.30-0.49

0.50+

*Significant at $p < .05$.

Note: The table shows pairwise correlations between the Insight survey and the District of Columbia Public Schools' total IMPACT score for the time period noted in the column heading. The first three columns have nonresponse weights applied, while the last three do not. The color intensity indicates the magnitude of the absolute value of the correlations, using the legend above. The Insight index is missing because weights cannot be estimated for the index.

Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2017/18 and 2018/19.

Table D14. Correlations between Panorama scales and principals' total School Leader IMPACT score with nonresponse weights applied (2017/18 to 2018/19)

Survey measure	Weighted both years	Weighted 2017/18	Weighted 2018/19	Unweighted both years	Unweighted 2017/18	Unweighted 2018/19
Perseverance	0.416*	0.507*	0.326*	0.414*	0.482*	0.352*
Rigorous Expectations	0.191*	0.215*	0.164	0.190*	0.169	0.213*
Professional Learning about SEL	0.289*	0.321*	0.253*	0.282*	0.291*	0.273*
Educating All Students	0.116	0.208*	0.005	0.070	0.115	0.024
Staff Engagement	0.285*	0.347*	0.220*	0.278*	0.310*	0.249*
Observations		107	108		107	108

Absolute value of the correlation:

0-0.09

0.10–0.29

0.30-0.49

0.50+

SEL is social-emotional learning.

*Significant at $p < .05$.

Note: The table shows pairwise correlations between the Panorama survey and the District of Columbia Public Schools' total School Leader IMPACT score for the time period noted in the column heading. The first three columns have nonresponse weights applied, while the last three do not. The color intensity indicates the magnitude of the absolute value of the correlations, using the legend above.

Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2017/18 and 2018/19.

Table D15. Correlations between Insight domains and principals' IMPACT subscores with nonresponse weights applied (2017/18 to 2018/19)

Survey measure	Weighted Student Outcome Goals score	Weighted Leadership Framework score	Leadership Framework score (cycle 1)	Leadership Framework score (cycle 2)
Weighted				
Academic Opportunity	0.206*	0.403*	0.534*	0.576*
Evaluation	0.093	0.246*	0.359*	0.342*
Instructional Planning for Student Growth	0.121	0.293*	0.353*	0.377*
Leadership	0.140*	0.290*	0.364*	0.375*
Learning Environment	0.198*	0.376*	0.464*	0.499*
Observation & Feedback	0.084	0.221*	0.287*	0.271*
Peer Culture	0.135*	0.336*	0.388*	0.423*
Professional Development	0.111	0.268*	0.328*	0.369*
School Operations	0.114	0.438*	0.465*	0.517*
Diversity, Equity, and Inclusion	0.109	0.374*	0.420*	0.431*
Family & Community Engagement	0.167*	0.356*	0.470*	0.490*
Unweighted				
Academic Opportunity	0.190*	0.396*	0.523*	0.547*
Evaluation	0.106	0.212*	0.334*	0.302*
Instructional Planning for Student Growth	0.126	0.271*	0.331*	0.352*
Leadership	0.136*	0.271*	0.348*	0.345*
Learning Environment	0.188*	0.354*	0.443*	0.466*
Observation & Feedback	0.089	0.180*	0.253*	0.232*
Peer Culture	0.127	0.312*	0.376*	0.407*
Professional Development	0.120	0.240*	0.306*	0.338*
School Operations	0.123	0.395*	0.432*	0.467*
Diversity, Equity, and Inclusion	0.121	0.346*	0.407*	0.398*
Family & Community Engagement	0.165*	0.331*	0.469*	0.462*

Absolute value of the correlation:

0-0.09

0.10–0.29

0.30-0.49

0.50+

*Significant at $p < .05$.

Note: The table shows pairwise correlations between the Insight domains and the District of Columbia Public Schools' School Leader IMPACT subscores for the time period noted in the column heading. The first section includes nonresponse weights when calculating school averages of each survey measure. The second section is unweighted. As described in box 1, the District of Columbia Public Schools' School Leader IMPACT final scores are an average of the weighted Student Outcome Goals score and the weighted Leadership Framework score. The weighted Leadership Framework score consists of two Leadership Framework scores from cycle 1 and cycle 2. The color intensity indicates the magnitude of the absolute value of the correlations, using the legend above. The Insight index is missing because weights cannot be estimated for the index.

Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table D16. Correlations between Panorama scales and principals' IMPACT subscores with nonresponse weights applied (2017/18 to 2018/19)

Survey measure	Weighted Student Outcome Goals score	Weighted Leadership Framework score	Leadership Framework score (cycle 1)	Leadership Framework score (cycle 2)
Weighted				
Perseverance	0.257*	0.393*	0.508*	0.560*
Rigorous Expectations	0.165*	0.099	0.142*	0.164*
Professional Learning about SEL	0.159*	0.305*	0.378*	0.398*
Educating All Students	0.123	0.018	0.031	-0.012
Staff Engagement	0.144*	0.324*	0.389*	0.385*
Unweighted				
Perseverance	0.263*	0.376*	0.485*	0.551*
Rigorous Expectations	0.186*	0.060	0.129	0.195*
Professional Learning about SEL	0.172*	0.270*	0.361*	0.377*
Educating All Students	0.113	-0.057	-0.039	-0.035
Staff Engagement	0.156*	0.291*	0.366*	0.373*

Absolute value of the correlation:

0-0.09	0.10–0.29	0.30-0.49	0.50+
--------	-----------	-----------	-------

SEL is social-emotional learning.

*Significant at $p < .05$.

Note: The table shows pairwise correlations between the Panorama scales and the District of Columbia Public Schools' School Leader IMPACT subscores for the time period noted in the column heading. The first section includes nonresponse weights when calculating school averages of each survey measure. The second section is unweighted. As described in box 1, the District of Columbia Public Schools' School Leader IMPACT final scores are an average of the weighted Student Outcome Goals score and the weighted Leadership Framework score. The weighted Leadership Framework score consists of two Leadership Framework scores from cycle 1 and cycle 2. The color intensity indicates the magnitude of the absolute value of the correlations, using the legend above. The Insight index is missing because weights cannot be estimated for the index.

Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2017/18 to 2018/19.

Table D17. Correlations between Insight domains and principals' Leadership Framework subscores with nonresponse weights applied (2018/19)

Survey measure	Instruction	Talent	School Culture	Operations	Family & Community	Personal Leadership
Weighted						
Academic Opportunity	0.532*	0.396*	0.470*	0.051	0.133	0.115
Evaluation	0.286*	0.172	0.318*	-0.107	0.091	0.037
Instructional Planning for Student Growth	0.260*	0.235*	0.314*	-0.058	0.103	0.135
Leadership	0.235*	0.160	0.301*	-0.058	0.146	0.128
Learning Environment	0.371*	0.248*	0.409*	0.031	0.205*	0.154
Observation & Feedback	0.211*	0.143	0.207*	-0.016	0.064	0.041
Peer Culture	0.317*	0.198*	0.357*	-0.007	0.124	0.118
Professional Development	0.278*	0.224*	0.298*	-0.055	0.158	0.148
School Operations	0.345*	0.219*	0.456*	0.077	0.306*	0.248*
Diversity, Equity, and Inclusion	0.293*	0.243*	0.386*	-0.011	0.195*	0.131
Family & Community Engagement	0.368*	0.293*	0.444*	-0.001	0.219*	0.129
Unweighted						
Academic Opportunity	0.514*	0.396*	0.441*	0.028	0.093	0.126
Evaluation	0.253*	0.156	0.288*	-0.042	0.043	0.028
Instructional Planning for Student Growth	0.230*	0.218*	0.260*	-0.018	0.073	0.120
Leadership	0.226*	0.154	0.269*	-0.025	0.113	0.126
Learning Environment	0.356*	0.240*	0.365*	0.027	0.170	0.160
Observation & Feedback	0.141	0.124	0.145	0.030	0.024	0.021
Peer Culture	0.323*	0.190	0.323*	0.008	0.105	0.139
Professional Development	0.238*	0.211*	0.259*	-0.021	0.129	0.130
School Operations	0.307*	0.186	0.429*	0.069	0.271*	0.221*
Diversity, Equity, and Inclusion	0.273*	0.225*	0.349*	0.010	0.182	0.119
Family & Community Engagement	0.359*	0.303*	0.460*	-0.008	0.237*	0.115
Observations	108	108	108	108	108	108

Absolute value of the correlation:

0-0.09 0.10–0.29 0.30-0.49 0.50+

*Significant at $p < .05$.

Note: The table shows pairwise correlations between the Insight survey domains and principals' Leadership Framework subscores, specifically the average of the subscore for each standard from cycle 1 and cycle 2. These detailed Leadership Framework component scores were only available for 2018/19. The color intensity indicates the magnitude of the absolute value of the correlations, using the legend above. The upper panel includes nonresponse weights when calculating the school averages for each survey measure, while the lower panel is unweighted. The Insight index is missing because weights cannot be estimated for the index.

Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2018/19.

Table D18. Correlations between Panorama scales and principals' Leadership Framework subscores with nonresponse weights applied (2018/19)

Survey measure	Instruction	Talent	School Culture	Operations	Family & Community	Personal Leadership
Weighted						
Perseverance	0.531*	0.347*	0.374*	0.135	0.037	0.096
Rigorous Expectations	0.122	0.016	0.111	0.105	-0.187	-0.091
Professional Learning about SEL	0.302*	0.259*	0.266*	0.173	0.137	0.114
Educating All Students	-0.172	-0.119	-0.050	-0.026	-0.094	-0.034
Staff Engagement	0.247*	0.258*	0.320*	0.051	0.110	0.167
Unweighted						
Perseverance	0.524*	0.357*	0.397*	0.124	0.028	0.083
Rigorous Expectations	0.160	0.042	0.181	0.111	-0.192*	-0.074
Professional Learning about SEL	0.295*	0.239*	0.264*	0.187	0.119	0.109
Educating All Students	-0.138	-0.104	-0.015	-0.022	-0.093	-0.021
Staff Engagement	0.250*	0.243*	0.325*	0.057	0.092	0.167
Observations	108	108	108	108	108	108

Absolute value of the correlation:

0-0.09	0.10-0.29	0.30-0.49	0.50+
--------	-----------	-----------	-------

SEL is social-emotional learning.

*Significant at $p < .05$.

Note: The table shows pairwise correlations between the Panorama scales and principals' Leadership Framework subscores, specifically the average of the subscore for each standard from cycle 1 and cycle 2. These detailed Leadership Framework component scores were only available for 2018/19. The color intensity indicates the magnitude of the absolute value of the correlations, using the legend above. The upper panel includes nonresponse weights when calculating the school averages for each survey measure, while the lower panel is unweighted.

Source: Authors' analyses based on administrative data provided by the District of Columbia Public Schools, 2018/19.

Appendix E. Principal effect interpretations

This appendix provides intuitive interpretations of the principal effects from figure 3 in the main report.

Table E1. Principal effect interpretations (2015/16 to 2018/19)

Measure	Principal effect standard deviation	Interpretation in percentile units: “Compared with the median principal in the district, a principal at the 90th percentile of the distribution of principal effects for this measure would move the median school’s score for this measure from the 50th percentile to the...”
Insight domains		
Index	0.47	73rd percentile
Academic Opportunity	0.40	70th percentile
Evaluation	0.47	73rd percentile
Instructional Planning	0.52	75th percentile
Leadership	0.61	78th percentile
Learning Environment	0.45	72nd percentile
Observation & Feedback	0.46	72nd percentile
Peer Culture	0.39	69th percentile
Professional Development	0.55	76th percentile
School Operations	0.45	72nd percentile
Student achievement		
*Percent proficient	0.07	54th percentile

*Percent proficient is the school average of the English language arts and math proficiency rates on the Partnership for Assessment of Readiness for College and Careers exam.

Note: The table takes the true standard deviation of the principal effects shown in figure 3 in the main report and provides an intuitive interpretation in percentile units.

Source: Authors’ analyses based on survey data provided by the District of Columbia Public Schools, 2015/16 to 2018/19.